

#ABCforJobs



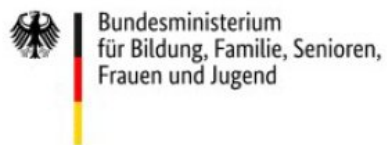
Evaluationsbericht zu den Serious Games „Escape The Plate“, „Sky Dash“, „Syllable Soup“ und „Laut-Manager“

Verfasser*innen: Geeske Kemper, Polona Caserman, Giulio Crocco, Benjamin Schnitzer, Stefan Göbel

8. August 2025



Gefördert vom:



Das Projekt #ABCforJobs

#ABCforJobs ist ein Modellprojekt, das von der GFFB gGmbH in Zusammenarbeit mit einem Konsortium aus 15 Kooperationspartner*innen koordiniert wurde. Ziel des Projekts war die Entwicklung und Erprobung schriftsprachlicher, arbeitsplatzorientierter Lerninhalte für gering literalisierte Erwachsene (glE) – sowohl in Beschäftigung als auch in Erwerbslosigkeit auf einem Alpha-Level 3 und Alpha-Level 4.

Der Schwerpunkt lag auf Branchen, die besonders stark von der Corona-Pandemie betroffen waren, wie dem Hotel- und Gaststättengewerbe (HoGa) sowie der Gepäckabfertigung am Frankfurter Flughafen (Gepa). Durch den Erwerb schriftsprachlicher und digitaler Grundkompetenzen sollten gering literalisierte Personen aus diesen Bereichen neue berufliche Perspektiven erhalten.

Evaluation der Serious Games

Im Rahmen der Entwicklung und Testphase der Serious Games wurde eine wissenschaftliche Evaluation durchgeführt, um Qualität, Nutzererlebnis, Gamedesign und Wirksamkeit zu bewerten. Ziel dieser Untersuchung war es, anhand von Befragungen, Beobachtungen, standardisierten Fragebögen und technischen Analysen fundierte Erkenntnisse zu Spielqualität, Gamedesign und Lernerfolg zu gewinnen und gezielt in die Weiterentwicklung einfließen zu lassen.

Projektlaufzeit: 01.11.2021 bis 31.10.2025

Evaluationszeitraum: 01/2024 bis 05/2025

Diese Publikation wurde im Rahmen der Nationalen Dekade für Alphabetisierung und Grundbildung 2016-2026 mit Mitteln des Bundesministeriums für Bildung, Familie, Senioren Frauen und Jugend unter dem Förderkennzeichen W-1505A-AOG gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor*innen.

Inhaltsverzeichnis

Das Projekt #ABCforJobs.....	2
Evaluation der Serious Games.....	2
1. Einführung.....	3
2. Escape the Plate (Escape Room).....	3
2.1 Einführung.....	3
2.2 Evaluation.....	4
2.2.1 Evaluationsaufbau.....	4
2.2.2 Durchführung.....	5
2.3 Evaluationsergebnisse und Schlussfolgerungen.....	6
2.3.1 Iteration 1.....	6
2.3.2 Iteration 2.....	8
3. Sky Dash - ein niedrigschwelliges Serious Game für mobile Endgeräte.....	9
3.1 Einführung.....	9
3.2 Methodik.....	10
3.3 Evaluationsergebnisse.....	10
3.4 Konklusion.....	11
4. Syllable Soup - ein Serious Virtual Reality Game.....	14
4.1 Einführung.....	14
4.2 Methodik.....	15
4.3 Evaluationsergebnisse.....	17
4.4 Konklusion.....	20
5. Laut Manager - ein Desktop-basiertes Serious Game zur Schriftsprachförderung.....	21
5.1 Einführung.....	21
5.2 Methodik.....	22
5.3 Evaluationsergebnisse.....	23
5.4 Konklusion.....	26
1. Feingranulare Adaptivität auf drei Ebenen.....	27
2. Multimodale, wertschätzende Rückmeldungen.....	27
3. Didaktische Erweiterung durch Mikro-Lernpfade.....	27
4. Blended-Learning-Szenario und Skalierung.....	27
6. Zusammenfassung der Evaluationsergebnisse.....	28
Literatur.....	29
7. Anhang.....	29
7.1 IMI-Fragebogen in einfacher Sprache.....	29
7.2 Fragebogen Sky Dash.....	29

1. Einführung

Über sechs Millionen erwachsene Menschen in Deutschland gelten als gering literalisiert. Gerade für diese Zielgruppe ist der Zugang zu klassischen Bildungsangeboten oft erschwert, sei es durch negative Schulerfahrungen, geringe Selbstwirksamkeit oder den schambehafteten Umgang mit dem sensiblen Thema der Literalität. An dieser Stelle können digitale Angebote, spielerische Formate mit niedrigschwelliger Zugänglichkeit und immersive Lernwelten, die durch direktes Feedback sowohl intrinsisch als auch extrinsisch motivierend wirken können, ihr Potenzial entfalten.

Vor diesem Hintergrund wurde im Rahmen von #ABCforJobs vier Serious Games „Escape the Plate“, „Sky Dash“, „Crazy Kitchen“ und „Laut Manager“ entwickelt. Dieser Bericht befasst sich mit der Evaluation dieser vier Serious Games. Zunächst wird zu jedem Spiel eine kurze Einführung gegeben und ein Überblick über die jeweilige Evaluation gegeben. Anschließend werden der Evaluationsaufbau und die Durchführung beschrieben. Darauf folgen die Darstellung und Diskussion der Ergebnisse. Am Ende jedes Abschnitts werden schließlich Schlussfolgerungen und Empfehlungen zur Weiterentwicklung des jeweiligen Spiels aufgeführt.

2. Escape the Plate (Escape Room)

2.1 Einführung

„Escape the Plate“ ist ein Lernspiel, das speziell für gering literalisierte Mitarbeitende und Interessierte im Bereich der Hotellerie und Gastronomie auf Alpha Level 3 und Alpha Level 4 im Rahmen des Projektes #ABCforJobs entwickelt wurde. Die Spielenden müssen drei verschiedene Aufgaben lösen, um aus einem Raum zu entkommen. Diese Aufgaben sind im Setting einer Hotelbar gestaltet.

„Escape the Plate“ wurde mit der Spiel-Engine Unity umgesetzt. Die Spielstruktur basiert auf einem digitalen Escape Room, in dem die Spieler*innen verschiedene Aufgaben lösen müssen, um eine Zahlenkombination zu erhalten. Jede dieser Aufgaben ist als interaktives Rätsel gestaltet – zum Beispiel zum Thema Silbentrennung, Rechtschreibung oder Lückentexte (siehe Abbildung 1).

Die Aufgaben und ihre Inhalte wurden so eingebaut, dass sie flexibel an zwei Schwierigkeitsstufen (Alpha-Level 3 und 4) angepasst werden können. Dafür wurde ein einfaches System entwickelt, das die passenden Inhalte je nach ausgewähltem Level im Spiel anzeigt. Das Spiel wurde so programmiert, dass es leicht zu bedienen ist und die Spieler*innen nach jeder gelösten Aufgabe direkt Rückmeldung bekommen. Ziel war es, die Lerninhalte möglichst verständlich und gleichzeitig spielerisch zu verpacken.

Ziel der Evaluation ist es zu überprüfen, ob das Spiel einen Lerneffekt erzielen kann und welche Rückmeldungen die Nutzer*innen zur Nutzbarkeit, Verständlichkeit und zum Spielspaß geben. Die Ergebnisse sollen einen ersten Einblick geben, ob das Spiel in seiner Form sinnvoll ist und an welchen Stellen es weiterentwickelt werden sollte. Da es sich um einen Prototypen handelt, stehen dabei vor allem qualitative Rückmeldungen im Vordergrund.

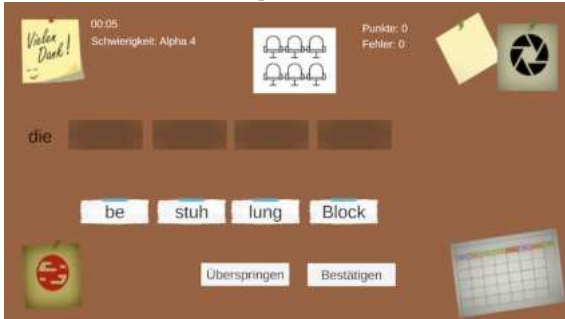
An der Evaluation nahmen nur einige Personen teil, insgesamt wurden zwei Iterationen durchgeführt (siehe Abschnitt 2.2). Die Ergebnisse sind daher nur als erste Einschätzung zu verstehen und statistisch nicht belastbar. Der Schwerpunkt liegt auf qualitativen Rückmeldungen, um Stärken und Schwächen des Prototyps zu identifizieren. Aussagen zur langfristigen Wirksamkeit sind nur eingeschränkt möglich, die Übertragbarkeit auf andere Zielgruppen ist begrenzt.



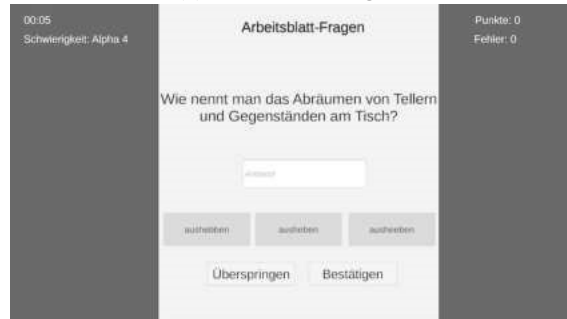
(a) Spielsicht



(b) Lückentext-Aufgabe



(c) Silbentrennung



(d) Rechtschreib-Aufgabe

Abbildung 1: Verschiedene Aufgaben und Ansichten aus Escape the Plate

2.2 Evaluation

Die beiden Iterationen fanden am 10.08.2023 bzw. im Zeitraum vom 11.03.2025 bis zum 18.03.2025 statt und unterschieden sich leicht im Ablauf, Aufbau sowie in der Durchführung. An den entsprechenden Stellen wird darauf eingegangen.

2.2.1 Evaluationsaufbau

Iteration 1 Die erste Iteration umfasste folgende Schritte:

- **Einleitung** (ca. 30 Minuten): Unterzeichnung der Einverständniserklärung und Ausfüllen eines demografischen Fragebogens.
- **Pre-Test** (ca. 20 Minuten)
- **Spielphase** (ca. 15 Minuten)
- **Post-Test** (ca. 20 Minuten)
- **Pause** (ca. 5 Minuten)
- **IMI-Fragebogen** (Intrinsic Motivation Inventory [CFSDT25], ca. 20 Minuten). Der Fragebogen wurde in einfacher Sprache übersetzt (siehe Anhang 7.1).

Insgesamt dauerte die erste Evaluation etwa 1 Stunde und 20 Minuten. Eine Übersicht ist in Abbildung 28 dargestellt.



Abbildung 2: Evaluationsaufbau Escape the Plate – Iteration 1

Iteration 2 Die zweite Iteration wurde aufgrund der Erfahrungen aus der ersten Iteration auf zwei Tage verteilt, um die Teilnehmer*innen nicht zu überfordern. Der Ablauf war wie folgt gegliedert:

- **Erster Tag:**

- Einleitung mit Einverständniserklärung und demografischem Fragebogen
- Pre-Test
- Pause (ca. 10 Minuten) – Spielphase (Aufgaben 1 und 2)

- **Zweiter Tag:**

- Spielphase (Aufgabe 3)
- IMI-Fragebogen
- Pause (ca. 10 Minuten)
- Post-Test

Eine Übersicht ist in Abbildung 3 dargestellt.



Abbildung 3: Evaluationsaufbau „Escape the Plate“ – Iteration 2

2.2.2 Durchführung

Vor Beginn der Evaluation wurden die Teilnehmer*innen über die Datenerhebung informiert und unterzeichneten eine Einverständniserklärung. Alle erhobenen Daten wurden anschließend anonymisiert ausgewertet.

Iteration 1 An der ersten Iteration nahmen insgesamt sechs Personen teil. Das Durchschnittsalter betrug 23,16 Jahre. Die Gruppe bestand aus fünf Teilnehmern und einer Teilnehmerin. Hinsichtlich des Schulabschlusses verfügte eine Person über einen Realschulabschluss in Deutschland, während fünf Personen einen Schulabschluss aus ihrem Heimatland ohne Anerkennung in Deutschland hatten. Die Evaluation dauerte insgesamt ca. zwei Stunden (9:30 bis 11:20 Uhr). Die meisten Teilnehmer*innen hatten einen Migrationshintergrund. Zudem zeigte sich, dass die lange Testdauer (insbesondere durch Pre- und Post-Test) sehr zeitintensiv war.

Überarbeitung des Spiels Basierend auf den Ergebnissen und Erfahrungen aus der ersten Iteration wurden Spiel und Studiendesign überarbeitet. Dazu gehörte die lokale Speicherung des Spiels, damit es an mehreren Tagen genutzt werden konnte, sowie die Verwaltung der Spieldaten ohne Serveranbindung. Um den Fortschritt der Teilnehmer*innen verfolgen zu können, wurde das Spiel an einen Learning Record Store (Learning Locker) angebunden. Zusätzlich wurde eine webbasierte Version des Spiels entwickelt, um mehr Flexibilität zu ermöglichen.

Iteration 2 Die zweite Iteration wurde anderthalb Jahre später durchgeführt. Dabei wurden bereits die Anpassungen aus der ersten Iteration umgesetzt. Insgesamt nahmen zwei Teilnehmer und eine Teilnehmerin teil. Das Durchschnittsalter lag bei 42,33 Jahren ($\pm 6,43$ Jahre). Alle Teilnehmer*innen hatten einen Migrationshintergrund. Die Evaluation wurde auf zwei Tage aufgeteilt, um die beschriebene Überforderung zu vermeiden. Diese Anpassung erfolgte aufgrund der Erfahrungen aus der ersten Iteration. Der erste Termin fand

am 11.03.2025 statt, der zweite eine Woche später am 18.03.2025. Trotz der Änderungen nahm eine kleine Gruppe teil, da es an der Größe der Gruppe lag, die das Trainingsangebot wahrnahm.

2.3 Evaluationsergebnisse und Schlussfolgerungen

Die Daten wurden deskriptiv ausgewertet, da die Stichprobengröße für weitergehende statistische Verfahren nicht ausreicht.

2.3.1 Iteration 1

Aus der ersten Evaluationsiteration ergeben sich folgende Ergebnisse und Schlussfolgerungen:

Wissenstest (Pre- und Post-Test) Die Ergebnisse des Wissenstests sind in Abbildung 4 dargestellt. Sie zeigen, dass die Leistungen im Post-Test in fast allen Kategorien schlechter ausfielen als im Pre-Test. Eine Ausnahme bildet die Kategorie Silbentrennung, in der sich die Werte leicht verbessert haben.

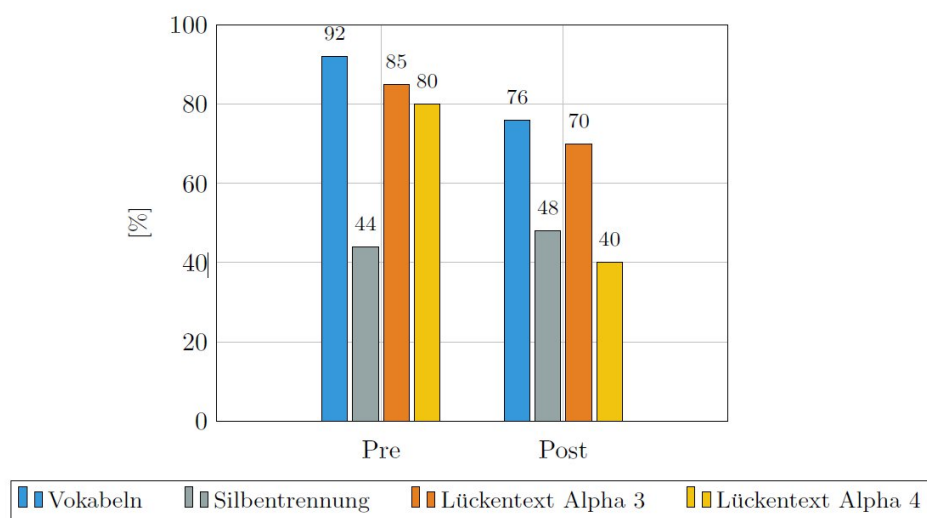


Abbildung 4: Ergebnisse Pre- und Post-Test aus Iteration 1

Game Part Die Ergebnisse des Game Parts sind in Abbildung 5 zu sehen. Die Auswertung des IMI-Fragebogens verdeutlicht, dass die Teilnehmer*innen generell Freude am Spiel hatten und es als wertvoll empfanden. Ihre eigene Kompetenz schätzten sie dagegen eher gering ein. Dies könnte auch an technischen Schwierigkeiten liegen, da einige Proband*innen Probleme mit der Steuerung hatten.

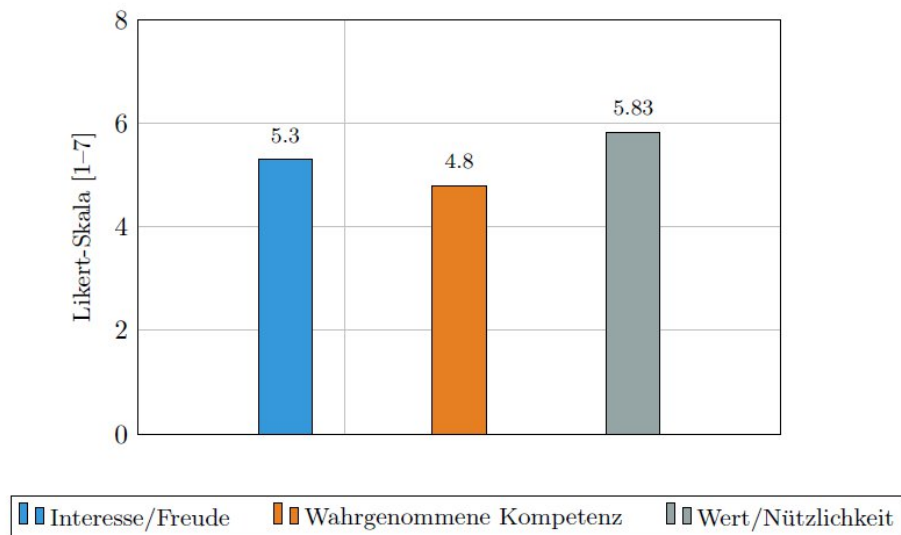


Abbildung 5: IMI-Fragebogen: Ergebnisse Game Part aus Iteration 2

Schlussfolgerungen Wissen Iteration 1 Die Verschlechterung der Ergebnisse im Post-Test lässt sich auf verschiedene Ursachen zurückführen:

- **Zufallseffekt:** Aufgrund der geringen Teilnehmer*innenzahl sind statistische Verzerrungen möglich.
- **Lerntransfer:** Der Transfer der im Spiel erworbenen Fähigkeiten auf den Wissenstest könnte nicht ausreichend gelungen sein.
- **Ermüdung und Stress:** Die lange Testdauer und die kurzen Abstände zwischen Pre- und Post-Test führten wahrscheinlich zu Konzentrationsverlust. Beobachtungen zeigten, dass einige Teilnehmer*innen nach der Evaluation erschöpft waren.
- **Lernmethode:** Die Aufgaben im Spiel könnten überarbeitet werden, um den Lernerfolg zu erhöhen.

Hauptproblem war vermutlich die geringe Aufmerksamkeitsspanne, was eine Überarbeitung der Testbedingungen notwendig machte.

Schlussfolgerungen Game Part Iteration 1 Die Ergebnisse des IMI-Fragebogens zeigen, dass die Teilnehmer*innen generell Spaß am Spiel hatten und es als wertvoll empfanden. Ihre eigene Kompetenz schätzten sie jedoch eher als gering ein, was möglicherweise auch an technischen Schwierigkeiten lag, da einige Teilnehmer*innen Probleme mit der Steuerung hatten.

Daraus ergeben sich folgende Überlegungen für die Anpassung des Game Designs:

- Die Erwartungen an die Grafik waren teilweise zu hoch, insbesondere bei Digital Natives. Der aktuelle Entwicklungsstand ist noch auf Prototyp-Niveau.
- Kürzere Minispiele und mehrere Levels können helfen, Überforderung zu vermeiden.
- Nicht alle Inhalte sollten auf einmal gespielt werden.
- Die Motivation kann durch Belohnungen – eventuell auch in der realen Welt – gesteigert werden.
- Kooperative oder kompetitive Elemente konnten das Spielerlebnis zusätzlich verbessern.

Um diese Punkte umzusetzen, wurde beschlossen, zukünftige Evaluationen in kürzere Einheiten aufzuteilen. Eine Speicherfunktion wurde implementiert, damit die Teilnehmer*innen das Spiel an mehreren Tagen spielen konnten und nicht alles auf einmal absolvieren müssen.

2.3.2 Iteration 2

Aus der zweiten Evaluationsiteration ergeben sich folgende Ergebnisse und Schlussfolgerungen.

Wissenstest (Pre- und Post-Test) Die Ergebnisse des Wissenstests sind in Abbildung 6 dargestellt. Sie zeigen, dass die Leistungen im Post-Test in fast allen Kategorien besser oder gleich ausfielen im Vergleich zum Pre-Test, mit Ausnahme der Kategorie Lückentext. Aufgrund der geringen Anzahl an Teilnehmer*innen lassen sich daraus keine belastbaren Aussagen ableiten. Auf den ersten Blick wirken die Ergebnisse jedoch positiver als in der Evaluation von 2023.

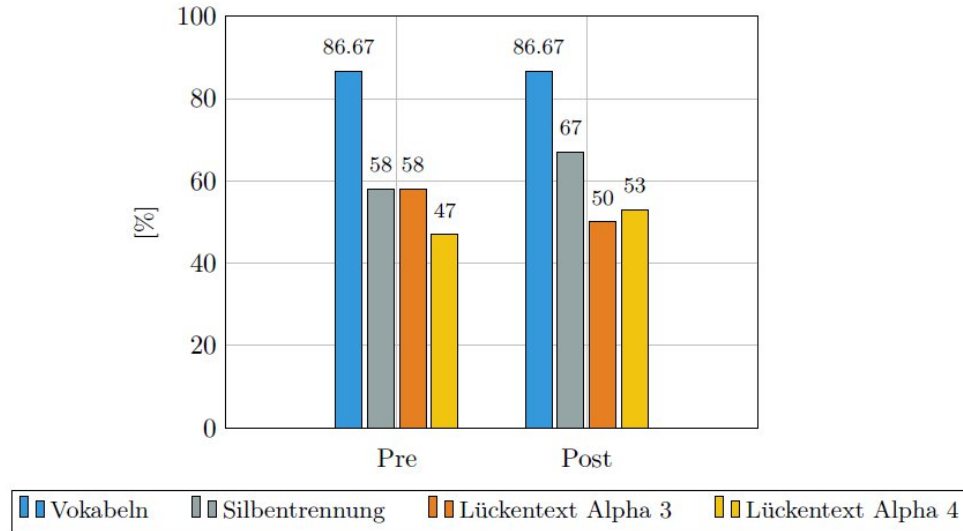


Abbildung 6: Ergebnisse Pre- und Post-Test aus Iteration 2

Game Part Die Ergebnisse des Game Parts sind in Abbildung 7 zu sehen. Sie ähneln den Ergebnissen aus dem Jahr 2023: Die Teilnehmer*innen hatten Spaß am Spiel und sehen darin einen potenziellen Nutzen für das Lernen.

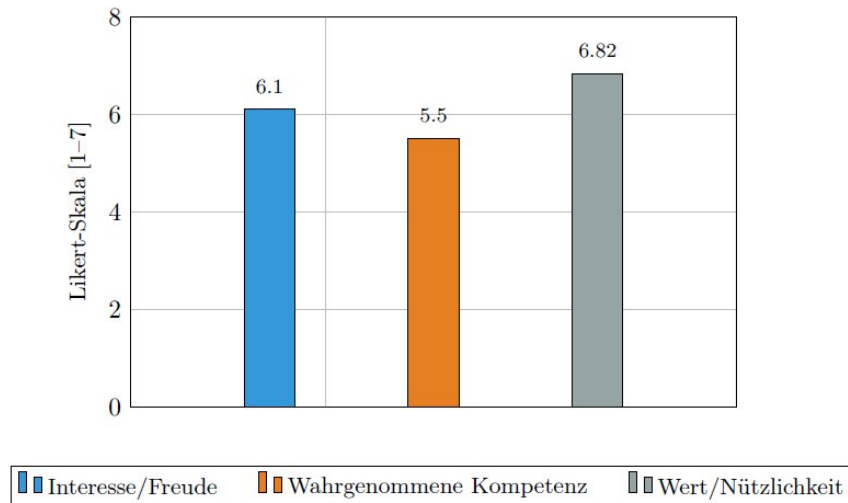


Abbildung 7: IMI-Fragebogen: Ergebnisse Game Part aus Iteration 2

Schlussfolgerungen Iteration 2 Die Ergebnisse der zweiten Iteration zeigen, dass das Spiel von den Teilnehmer*innen als motivierend und wertvoll wahrgenommen wird und Potenzial hat. Aufgrund der

vorhandenen kleinen Gruppengrößen konnten nicht mehr Teilnehmer*innen für die Evaluation gewonnen werden, was zu einer kleinen Stichprobe führte. Daher wurde „Escape the Plate“ in allen Kursen gespielt, sobald es nach der Anpassung einsatzbereit war.

3. Sky Dash - ein niedrigschwelliges Serious Game für mobile Endgeräte



Abbildung 8: Ausschnitte aus dem Spiel „Sky Dash“

3.1 Einführung

Im Rahmen der Zusammenarbeit im Projekt #ABCforJobs mit den Projektpartnern Fraport Ground Services (FRA) und Korion Interactive GmbH (KOR) wurde mit „Sky Dash“ ein Serious Game für mobile Endgeräte sowie den Browser (vgl. Abbildung 8) entwickelt und evaluiert. Ziel war es, Mitarbeitende eines internationalen Flughafens bei der Aneignung und Wiederholung von Flughafenabkürzungen (IATA 3-Letter Codes) zu unterstützen – einem typischerweise repetitiven und monotonen Lerninhalt, der der Zielgruppe gering literalisierter Erwachsener in

Bezug auf Lesegeschwindig- und Flüssigkeit sowie kognitiver Überlastung Probleme bereiten kann.

Das Spiel richtet sich insbesondere an Beschäftigte in der Gepäckabfertigung, für die ein effizientes und langfristig abrufbares Faktenwissen essenziell ist. Die zentralen Fragestellungen der Evaluation waren: (1) ob sich mit einem niedrigschwelligem, mobil verfügbarem Spiel ein messbarer Lerneffekt erzielen lässt, und (2) wie das subjektive Nutzererlebnis im Vergleich zu klassischen, nicht-gamifizierten Lernmethoden ausfällt.

„Sky Dash“ ist ein Lernspiel im Stile eines Endless Runners; Spielende steuern ein Flugzeug, das passend zu einem angegebenen Zielflughafen (ausgeschrieben als Wort), die entsprechenden Flughafenabkürzungen einsammeln muss. Für jede richtige Auswahl erhalten Spielende direktes Feedback und Punkte, mit denen sich nach und nach neue Welten und Flugzeuge freischalten lassen.

Neben der gezielten Förderung der Lesegeschwindigkeit werden weitere kognitive und wahrneh-

mungsbezogene Kompetenzen sowie sensomotorische Fähigkeiten und Reaktionszeit trainiert; alles zentrale Fähigkeiten im Alltag in der Gepäckabfertigung. Außerdem vermittelt das Serious Game Fachwissen rund um die international gebräuchlichen Drei-Buchstaben-Flughafencodes. Zentrale Elemente und Mechaniken des Spiels sind ein Punktesystem, ein Zeitlimit und Belohnungen in Form von Badges.

3.2 Methodik

Zur Beantwortung der Forschungsfragen wurde eine empirische Vergleichsstudie mit 79 Teilnehmer*innen (alle Personen standen zum Zeitpunkt der Evaluation am Anfang der Ausbildung zum*zur Gepäckabfertiger*in) durchgeführt, die in zwei Gruppen aufgeteilt wurden. Gruppe A (n = 39) arbeitete mit einem passivem Lernmedium (Video), das die zu lernenden IATA Codes (insgesamt 23 unterschiedliche Drei-Buchstaben-Codes) vermittelte; Gruppe B (n = 40) lernte mit Hilfe des interaktiven Lernmediums „Sky Dash“ die exakt selben Inhalte (insgesamt 23 unterschiedliche Drei-Buchstaben-Codes). Gruppe B erhielt ein zusätzliches Pre-Training, um sich an die Spielsteuerung zu gewöhnen, in dem abweichende IATA Codes eingesetzt wurden. Beide Gruppen wurden jeweils vor der Lernphase (Pre-Test), unmittelbar danach (Post-Test) und nach einer gewissen Verzögerung (Memory-Test) getestet. Gemessen wurde der Wissenserwerb anhand der Differenz aus korrekten Antworten aus dem Post- und Pre-Test sowie Memory- und Pre-Test sowie die subjektive Nutzererfahrung mithilfe etablierter Skalen zu Spielspaß, Flow, Feedbackwahrnehmung, wahrgenommener Selbstwirksamkeit, Fokus und Gesamtzufriedenheit. Zusätzlich fanden Fokusgruppengespräche statt, um qualitative Rückmeldungen zu erhalten. Die Evaluationen fanden im Zeitraum zwischen Juli 2023 und April 2024 statt.

Der Aufbau der Studie ist in Abbildung 9 dargestellt. Der verwendete Fragebogen ist im Anhang unter Abschnitt 7.2 zu finden.

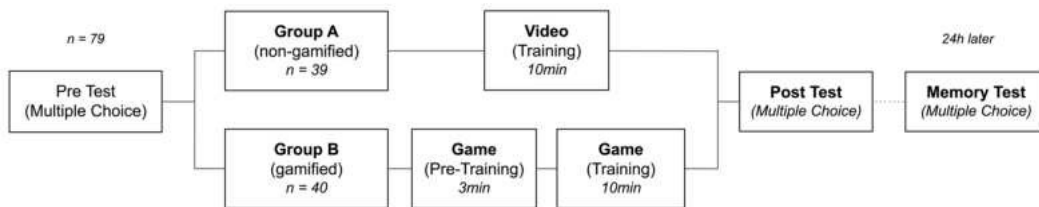


Abbildung 9: Studiendesign

3.3 Evaluationsergebnisse

Zur statistischen Auswertung wurden aufgrund nichtexistierender Normalverteilungen der Daten FriedmanTests zur Überprüfung der Messwiederholungen sowie Wilcoxon-Tests inklusive Bonferroni-Korrektur als Post-Hoc Vergleiche (Signifikanzniveau Alpha = 0,05) angewandt. Mann-Whitney-U-Tests wurden für Vergleiche zwischen den beiden unabhängigen Gruppen verwendet.

Die Ergebnisse zeigen, dass beide Gruppen signifikante Lernfortschritte zwischen dem Pre- und dem Post-Test sowie zwischen dem Pre- und dem Memory-Test erzielten. Dies weist darauf hin, dass sowohl das passive Lernmedium als auch das interaktive Spielformat grundsätzlich geeignet waren, den Zielinhalt effektiv zu vermitteln. Allerdings schnitt die nicht-gamifizierte Gruppe im direkten Vergleich signifikant besser ab. Diese Überlegenheit zeigte sich sowohl in der unmittelbaren Wissensabfrage nach dem Lernen als auch im verzögerten Erinnerungstest (vgl. Abbildung 10), wobei die Effektstärken in einem hohen Bereich lagen ($r \geq 0,55$). Darüber hinaus verdeutlicht Abbildung 11 den reinen Wissenszuwachs im Post- und Memory-Test im Vergleich zum Pre-Test. Hier wird der Unterschied zwischen den Gruppen noch einmal deutlich: Die nicht-gamifizierte-Gruppe erzielte einen signifikant höheren Wissenszuwachs.

Trotz dieses Unterschieds konnte auch die Spielgruppe ihr Wissen signifikant steigern. Dies ist insofern bemerkenswert, als dass die Lernumgebung hier spielerisch und weniger fokussiert auf explizites Memorieren

gestaltet war. Die Ergebnisse lassen darauf schließen, dass ein niedrigschwelliges Game wie „Sky Dash“ einen echten Wissenstransfer bewirken kann, auch wenn es nicht die maximale Effizienz klassischer Formate erreicht.

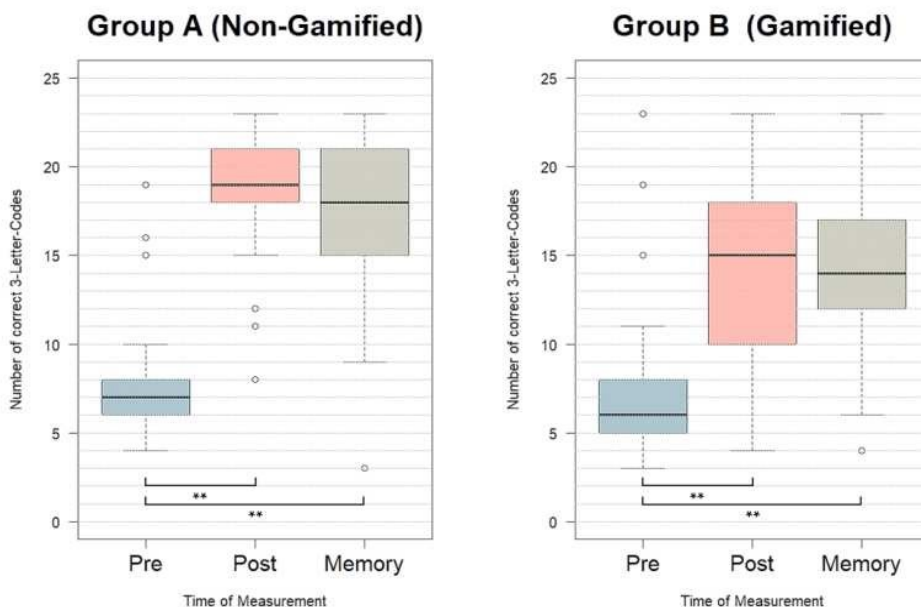


Abbildung 10: Vergleich der Lernleistung zwischen Kontrollgruppe (Non-Gamified) und Spielgruppe (Gamified) und über drei Testzeitpunkte

Bezüglich der User Experience zeigten sich ebenfalls interessante Unterschiede (vgl. Abbildung 12). Der Spielgruppe machte die Beschäftigung mit dem Lernstoff signifikant mehr Spaß als der Vergleichsgruppe, was durch moderate Effektstärken unterstrichen wurde. Weitere Dimensionen wie Flow-Erleben, Fokus, Feedbackwahrnehmung oder subjektive Selbstwirksamkeit unterschieden sich hingegen nicht signifikant zwischen den Gruppen. Die Rückmeldungen aus den Fokusgruppen darauf hin, dass „Sky Dash“ als unterhaltsam, ansprechend und alltagstauglich wahrgenommen wurde. Mehrere Teilnehmende gaben an, das Spiel gerne auch in der Freizeit oder auf dem Weg zur Arbeit einsetzen zu wollen.

Die Video-Gruppe schätzte dagegen die Klarheit der Präsentation und den Fokus auf Inhalte, äußerte jedoch auch Verbesserungsvorschläge - beispielsweise die Integration von Musik oder Farbgebung zur besseren Rezeption. Ein bemerkenswerter Aspekt war, dass die Video-Gruppe trotz des geringeren Spaßfaktors ihre eigene Lernleistung sehr positiv wahrnahm, was sich in hohen Selbstwirksamkeitswerten niederschlug. Eine detaillierte Darstellung der Ergebnisse ist in der dazugehörigen Publikation¹ zu finden.

3.4 Konklusion

Die Evaluation von „Sky Dash“ liefert differenzierte, praxisrelevante Ergebnisse. Das Serious Mobile Game erreicht sein Ziel, Lerninhalte auf niedrigschwellige und motivierende Weise zu vermitteln. Besonders für Inhalte mit hohem Wiederholungsbedarf und geringer emotionaler Bindung, wie es bei Flughafencodes der Fall ist, kann ein spielerischer Ansatz dazu beitragen, die Beschäftigung mit dem Lernstoff attraktiver zu gestalten. Die höhere Motivation und der signifikante Spaßfaktor bei der Spielgruppe sind Indikatoren für ein erhöhtes Potenzial zur langfristigen Lernbindung. Gerade in arbeitsintensiven Berufen mit geringer Lernzeit im Alltag kann dies einen entscheidenden Vorteil darstellen.

Allerdings zeigt sich auch, dass ein zu starker Fokus auf spielerische Elemente die kognitive Belastung erhöhen kann. Die Interaktivität und das permanente Feedback innerhalb des Spiels beanspruchen mentale Ressourcen,

¹ B. Schnitzer, P. Caserman, O. Korn, Sky Dash: Evaluating the Effects of a Serious Low-Threshold Mobile Game on Learning Efficacy and User Experience in a Repetitive Learning Task, in: J.L. Plass, X. Ochoa (Hg.), Serious Games. JCSG 2024, Lecture Notes in Computer Science, Bd. 15259, Cham 2025, https://doi.org/10.1007/978-3-031-74138-8_16.

die andernfalls vollständig für den Lerninhalt zur Verfügung stehen würden. Hieraus ergibt sich die Notwendigkeit, Spielmechaniken gezielt zu gestalten, um eine Balance zwischen kognitiver Aktivierung und Überforderung zu gewährleisten, gerade für die Zielgruppe gering literalisierter Erwachsener.

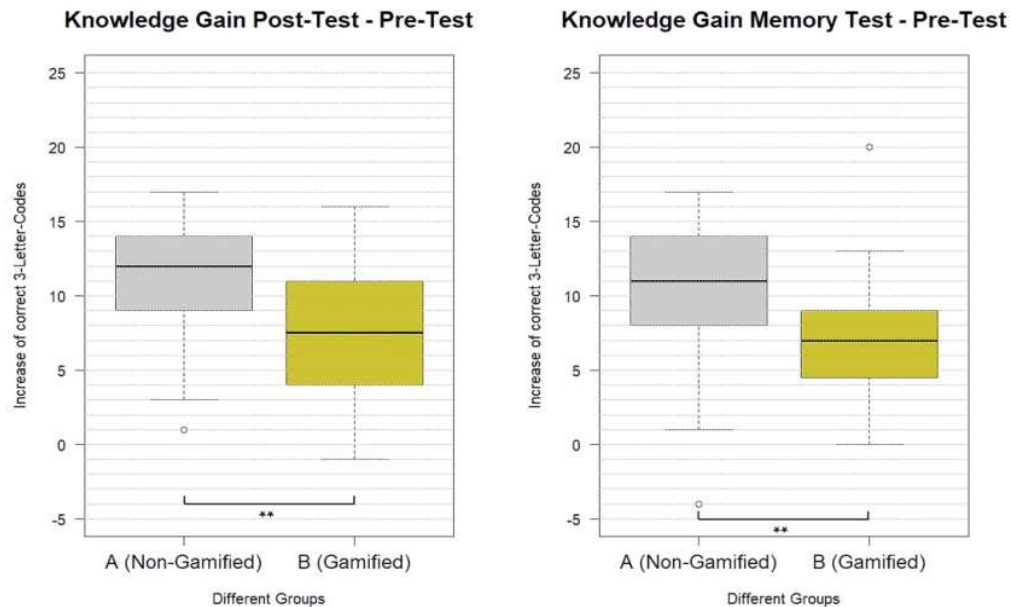


Abbildung 11: Wissenszuwachs: Vergleich der Gruppen im Post- und Memory-Test (Differenz zum Pre-Test)

Ein weiterer positiver Aspekt ist der mobile Zugang zum Spiel. Die Möglichkeit, Lernzeiten flexibel und ortsunabhängig zu nutzen, entspricht den Anforderungen moderner Arbeitsumfelder und wurde von den Teilnehmer*innen explizit als Vorteil hervorgehoben. Die vorliegende Studie bestätigt, dass mit „Sky Dash“ ein funktionaler Prototyp für mobiles, spielerisches Lernen entwickelt worden ist. Der Wissenserwerb ist nachweisbar und der Spielspaß hoch. Für den Praxiseinsatz und eine Skalierung empfehlen sich jedoch gezielte Weiterentwicklungen:

- Erstens sollte das Gameplay überarbeitet werden, um die kognitive Belastung zu verringern. Eine Reduktion der Ablenkung durch optische Reize und eine stärkere Konzentration auf den Lerninhalt, etwa durch adaptive Spielmechaniken oder gezielte Wiederholung bereits gelernter Codes, konnte die Effektivität deutlich steigern.
- Zweitens sollten weitere Untersuchungen zu nachhaltigem Wissenserhalt durchgeführt werden. Eine Langzeitstudie z. B. mit Tests nach vier oder acht Wochen konnte die tatsächliche Retention und damit die Alltagstauglichkeit noch besser bewerten.
- Drittens lässt sich „Sky Dash“ durch seine modulare Architektur leicht auf andere Themengebiete übertragen. Neben Flughafencodes wären auch Gefahrgutklassen, Sicherheitsvorgaben oder Standardprozesse geeignete Anwendungsfelder. Eine systematische Erweiterung der Inhalte konnte den Nutzen des Spiels in der betrieblichen Aus- und Weiterbildung erheblich steigern.

Darüber hinaus empfiehlt sich ein hybrider Lernansatz, bei dem klassisches und spielerisches Lernen kombiniert werden. So konnten beispielsweise kurze Videoeinheiten zur Einführung in ein Thema dienen, gefolgt von spielerischen Übungseinheiten mit „Sky Dash“. Dies würde eine effektive Kombination aus didaktischer Klarheit und motivationaler Bindung schaffen. Abschließend kann festgestellt werden, dass „Sky Dash“ ein vielversprechender Ansatz ist, um monotone, aber notwendige Lerninhalte auf motivierende Weise zu vermitteln. Trotz gewisser Einschränkungen in der reinen Lernleistung im Vergleich zur klassischen Methode spricht insbesondere der hohe Spaßfaktor für eine Integration solcher spielerischer Formate in Weiterbildungsmaßnahmen - auch bei der Zielgruppe gering literalisierter Erwachsener. Gerade unter dem Gesichtspunkt der Lernmotivation und des selbstgesteuerten Lernens bietet „Sky Dash“ einen wertvollen Beitrag zu modernen Lernumgebungen in der Arbeitswelt.

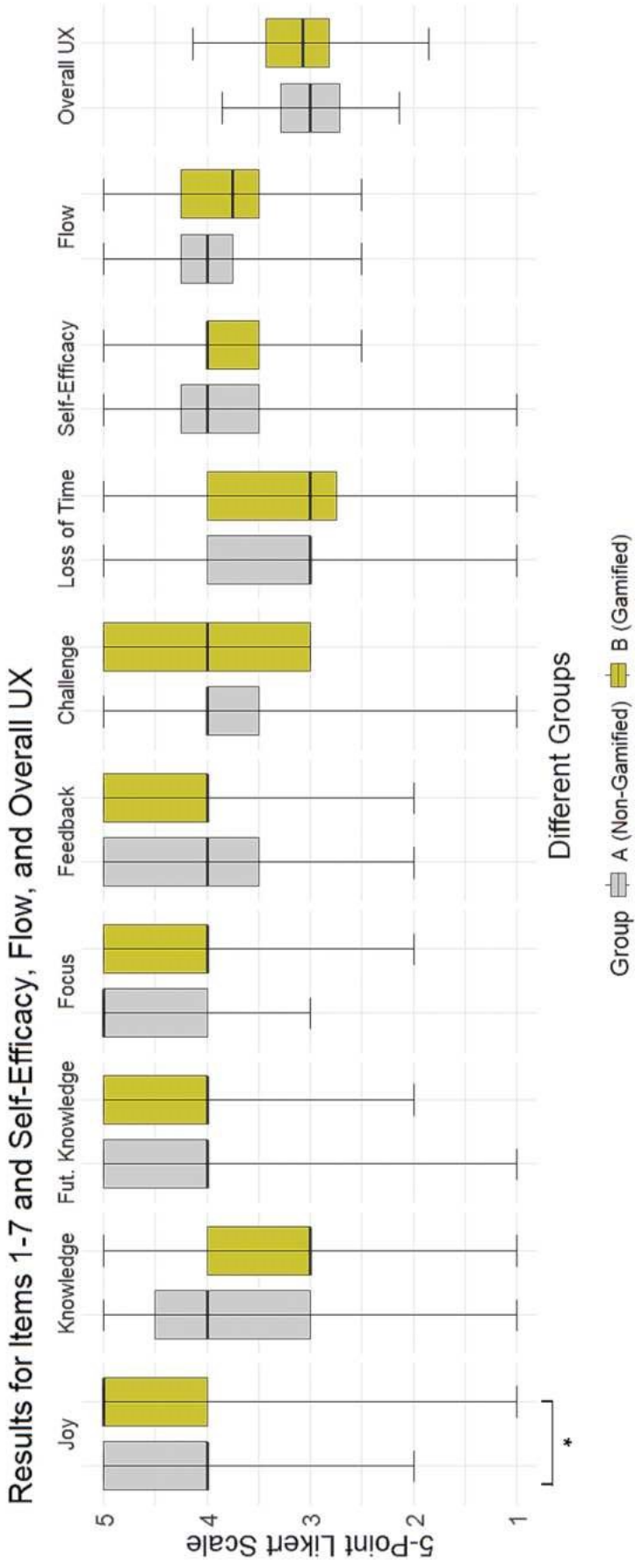


Abbildung 12: Selbsteinschätzung zu Motivation, Flow und Spielspaß (Vergleich beider Gruppen)

4. Syllable Soup - ein Serious Virtual Reality Game

4.1 Einführung



Abbildung 13: Teilnehmende spielen Syllable Soup

Ziel des im Projekt #ABCforJobs entwickelten Serious Games „Syllable Soup“ war es, die phonologische Bewusstheit, Worterkennung und Leseflüssigkeit gering literalisierter Erwachsener in der Hotel- und Gastronomiebranche durch ein vollimmersives, bewegungsbasiertes Lernspiel (Embodied Learning) zu fördern. Das Spiel ist ein sogenanntes Serious Virtual Reality Game (SVRG), das vollständig in einer virtuellen Küche stattfindet und mittels sog. Head-Mounted Display erlebt werden kann (vgl. Abbildung 13). Die Spielenden übernehmen die Rolle von Köch*innen und müssen die perfekte „Silbensuppe“ kochen. Dazu müssen sie umherfliegende Silben fangen und korrekt in Kochtöpfe sortieren, um so vollständige Wörter zu bilden.

Kernmerkmale des Spiels sind:

- **Embodiment:** Die Interaktion erfolgt physisch per Handbewegungen (Controller oder Leap Motion Kamera Sensor) und Fußbewegungen (mittels Fuß-Tracker).
- **Multisensorisches Feedback:** Audiovisuelle Rückmeldungen unterstützen den Lernprozess (grünes Häkchen bei korrekter Silbenfolge, Audioausgabe bei vollständigem Wort, virtuelle Belohnung durch Restaurant-Sterne).
- **Spielerische Mechaniken:** Elemente wie „heiße Topfdeckel“ mit Zeitlimit, Kakerlaken als Storelemente (die mit virtuellen Schuhen zertreten werden müssen) und ein Belohnungssystem steigern Motivation, Einbindung und Interaktion.
- **Niedrigschwellige Konzeption:** Die Spielsteuerung ist auf einfache Bewegungen reduziert, Inhalte stammen aus der Lebenswelt der Zielgruppe und alltagsnahen Arbeitsfeldern (Gastronomie, Service), das Interface ist bewusst reduziert und für die Zielgruppe visuell ansprechend gestaltet.

Gegenstand der Evaluation war die Frage, ob „Syllable Soup“ als immersives Spiel für die Zielgruppe

eine geeignete Lernumgebung darstellt und wie es sich im Vergleich zur Desktop-Version (am PC gespielt) desselben Spiels hinsichtlich Lernleistung, Usability und Nutzer*inerleben verhält. Dabei wurde insbesondere auch der Frage nachgegangen, ob sich durch VR und der damit einhergehenden stärkeren Immersion auch ein höherer Grad an Konzentration und zielgerichteterem Lernen erzielen lässt; bei gleichzeitigem Verzicht auf traditionelle, textbasierte Methoden.

Die Evaluation verfolgte zwei zentrale Fragestellungen: Zum einen sollte untersucht werden, ob SVRGs als motivierendes Werkzeug für gering literarisierte Menschen wertvoll sind. Zum anderen sollte überprüft werden, ob das SVRG als ergänzendes Lehrangebot den Lernprozess dieser Zielgruppe zusätzlich unterstützen können. Darüber hinaus wurde der Vergleich zwischen VR- und der PC-Version betrachtet, um mögliche Unterschiede in der Wirkung und Nutzungserfahrung zu ermitteln.

Die Evaluation wurde iterativ durchgeführt: Die erste Iteration (Pilotstudie) fand im vierten Quartal 2024 statt, die zweite im ersten Quartal 2025. Ergebnisse und Erkenntnisse aus der ersten Studie wurden dabei genutzt, um das Spiel und das Interaktionsdesign für die zweite Runde an die Bedarfe der Zielgruppe anzupassen.

4.2 Methodik

Die Studie folgte einem Gegenbalance-Within-Participants-Design, bei dem Erwachsene, die zuvor testbasiert als gering literalisiert eingestuft wurden, zwei Versionen des Spiels durchliefen: eine PC-gestützte, niedrig-immersive Variante und eine vollständig immersive VR-Version, erlebt durch ein HMD (*HTC Vive Pro 2*). Das Spielkonzept sah vor, Silben in der richtigen Reihenfolge in Töpfe zu sortieren, wobei visuelle, auditive und haptische Rückmeldungen zur Unterstützung gegeben wurden. In der VR-Variante erfolgte die Interaktion über Körperbewegungen, inklusive Fuß-Tracking und Objektmanipulation per Controller.

Die Teilnehmer*innen spielten je 10 Minuten pro Bedingung, mit Pre-, Post- und Memory-Tests zur Erhebung der Lernleistung (Silbenkombinationen), ergänzt durch standardisierte Fragebögen zu Usability (*SUS*), Flow (*FSS*) und User Engagement (*UES-SF*). Zusätzlich wurden In-Game-Leistungsmetriken (Fehlerrate, Reaktionszeit, Punktzahl) erfasst sowie qualitative Rückmeldungen und Videobeobachtungen analysiert.

Im Detail deckten die eingesetzten Fragebogen folgende Dimensionen ab:

- **Spaß, Fähigkeitsverbesserung, Nützlichkeit und Präsenz:** Die Teilnehmer*innen bewerteten, wie viel Spaß das Spiel gemacht hat, ob sie es als hilfreich für ihre Lese- und Schreibfähigkeiten einschätzten, ob sie es insgesamt als nützlich empfanden und wie stark sie sich als Teil der virtuellen Umgebung fühlten. Dabei zeigte sich in den Rückmeldungen zum Beispiel, dass die VR-Version als unterhaltsamer erlebt wurde, während der PC teilweise als besser geeignet für den Kompetenzerwerb eingeschätzt wurde.
- **User Engagement (O'Brien et al., 2018) [OCH18]:** Dieser Fragebogen erfasste, wie wertvoll die Nutzenden das Serious Game insgesamt einschätzten – abhängig davon, wie sehr sie mit dem Spiel interagierten oder sich engagierten. Abgefragt wurden u.a. gezielte Aufmerksamkeit, wahrgenommene Benutzerfreundlichkeit, ästhetischer Reiz und der Belohnungsfaktor.
- **System Usability (Bangor et al., 2009) [BKM09]:** Zur Bewertung der Benutzerfreundlichkeit kam ein etablierter Usability-Fragebogen zum Einsatz. Erste Ergebnisse deuten darauf hin, dass die Teilnehmer*innen die System Usability eher verhalten einschätzten - möglicherweise auch aufgrund einer geringen Technologieaffinität und einer fehlenden Eingewöhnungsphase für die Spielmechaniken.
- **Flow (Engeser und Rheinberg, 2008) [ER08]:** Der Flow-Fragebogen prüfte, ob die Teilnehmer*innen beim Spielen einen Zustand optimaler Leistung erlebten, ohne Über- oder Unterforderung.
- **VR-spezifische Fragen:** Zusätzlich wurden in der VR-Bedingung Aspekte wie Embodiment (Selbstwahrnehmung im virtuellen Körper mit den Dimensionen Ownership und Agency) sowie Cybersickness erfasst. Diese Items wurden auf einer 7-Punkte-Likert-Skala bewertet.

Nach der Pilotstudie mit sieben Teilnehmer*innen wurde das Spiel angepasst und u.a. Eingabemodalitäten vereinfacht und potenziell überfordernde Spielelemente wie die „Kakerlaken-Gegner“ entschärft bzw. positiv konnotiert.

Das Studiendesign blieb in beiden Iterationen unverändert. Eine Übersicht des Studiendesigns ist in Abbildung 14 dargestellt.

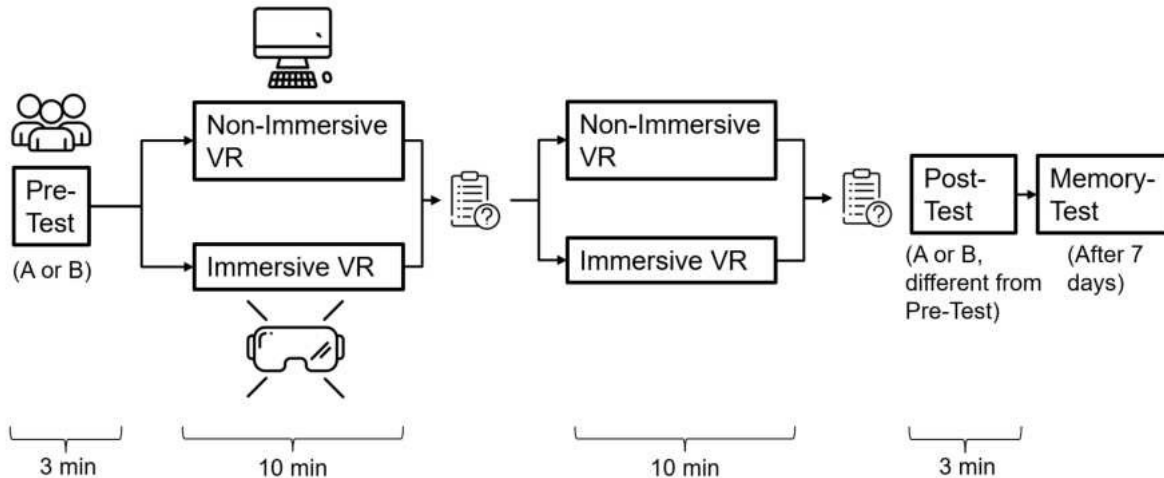


Abbildung 14: Studienaufbau Crazy Kitchen

Iteration 1 (Pilotstudie) An der ersten Iteration nahmen insgesamt sieben Teilnehmer*innen teil. Alle Teilnehmer*innen hatten einen Migrationshintergrund und insgesamt eine eher geringe Technologieaffinität, drei Teilnehmer*innen nutzen regelmäßig einen PC, zwei einmal pro Woche und zwei gaben an, nie einen PC zu nutzen. Keiner der Teilnehmer*innen hatte Vorerfahrungen mit Virtual Reality.

Sechs Teilnehmer*innen gaben an weiblich zu sein und eine Person männlich. Das Durchschnittsalter lag bei 43,43 Jahren ($\pm 6,95$ Jahre; zwischen 32 und 53 Jahren).

Anpassungen im Game Design Basierend auf den Erkenntnissen der ersten Iteration wurden mehrere Verbesserungsvorschläge abgeleitet (vgl. Abbildung 15):

- Akustisches Feedback und Sounddesign des Spiels sollte weniger von der eigentlichen Aufgabe ablenken und ausschließlich positiv konnotiert werden.
- Gegner (Kakerlaken) sorgen nicht für Minuspunkte, das Beseitigen kann jedoch zu Bonuspunkten (Streak) führen.
- Statt Leap Motion sollten VR-Controller eingesetzt werden, um die Steuerung zu erleichtern und Ausfälle des Trackings zu vermeiden.
- Einbindung der Füße durch weitere Sensorik, um das Spielerlebnis noch immersiver zu gestalten.
- Während die Silben mit den Händen gefangen werden, sollten Gegner (Kakerlaken) ausschließlich durch Treten der Füße beseitigt werden. Somit soll eine ganzheitliche Hand-Fuß-Koordination erreicht werden.

Iteration 2 Die zweite Iteration erfolgte mit einer ähnlichen Stichprobengröße, die Gruppe wies eine große Altersspanne auf. Alle Teilnehmer*innen hatten einen Migrationshintergrund und insgesamt eine geringe Affinität zu digitalen Technologien.

Es nahmen neun Personen teil (acht weiblich, eine männlich). Das Durchschnittsalter lag bei 41,56 Jahren ($\pm 8,65$ Jahre). Alle Teilnehmer*innen wurden im Ausland geboren.

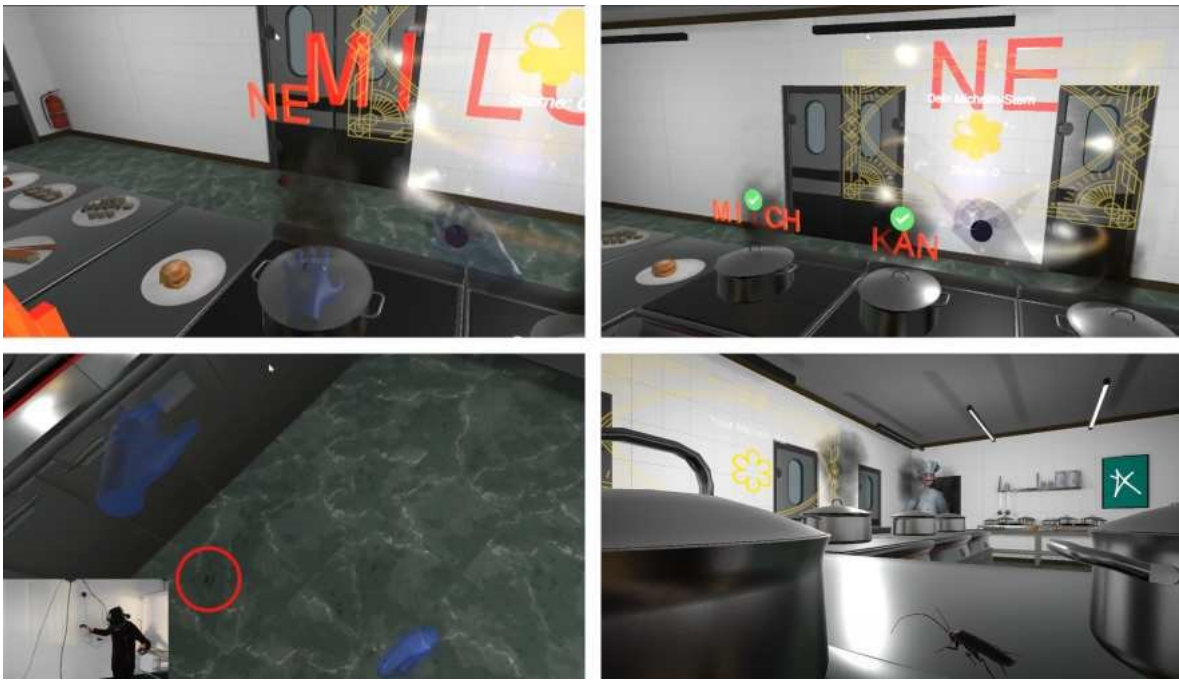


Abbildung 15: Gameplay mit angepasster Interaktionstechnik und verbesserten Feedbackmechanismen.

4.3 Evaluationsergebnisse

Iteration 1 (Pilotstudie) In der ersten Iteration wurden nur geringe Leistungsverbesserungen erzielt (siehe Abbildung 16), was vor allem auf eine zu hohe Belastung durch die Anzahl an Fragen und Bedingungen sowie auf Ermüdungseffekte zurückzuführen ist. Gleichzeitig gaben die Teilnehmer*innen an, dass sie das Spielprinzip motivierend fanden. Hervorzuheben ist hierbei besonders der erhöhte Spaßfaktor der VR Variante im Vergleich zur PC Version. Insgesamt zeigten die Ergebnisse aber auch, dass Verbesserungsbedarf bei Interaktionstechnik und beim Feedback bestand (siehe Abbildung 17).

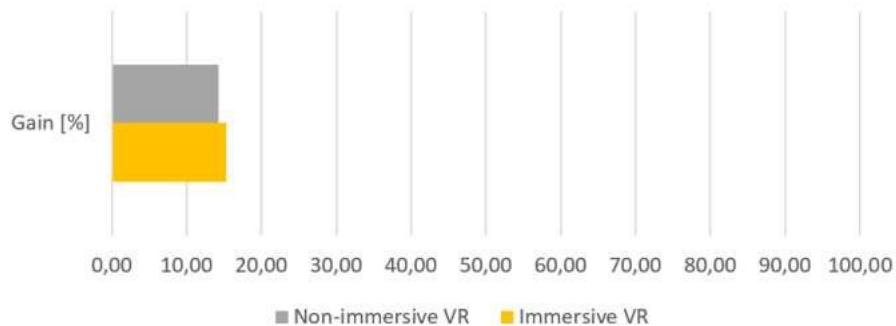


Abbildung 16: Wissenszuwachs zwischen Pre- und Post-Test in Iteration 1

Iteration 2 Die Ergebnisse der Studie lieferten differenzierte Einblicke in die Wirkungen des SVRG „Syllable Soup“ auf Lernleistung, Spielverhalten und Nutzer*inerleben bei gering literalisierten Erwachsenen. Die Teilnehmer*innen (n = 9) durchliefen sowohl eine low-immersive PC-Version als auch eine vollständig immersive VR-Version des Spiels, wobei Pre-, Post- und Memory-Tests sowie Fragebögen und In-Game-Metriken ausgewertet wurden.

Lernleistung: Beide Spielvarianten führten zu Lernfortschritten, jedoch in unterschiedlichem Ausmaß und mit abweichender zeitlicher Dynamik (vgl. Abbildung 18). Für die PC-Version zeigte ein Friedman-Test signifikante Unterschiede zwischen Pre-, Post- und Memory-Test ($\chi^2 = 10.56, p < .01$).

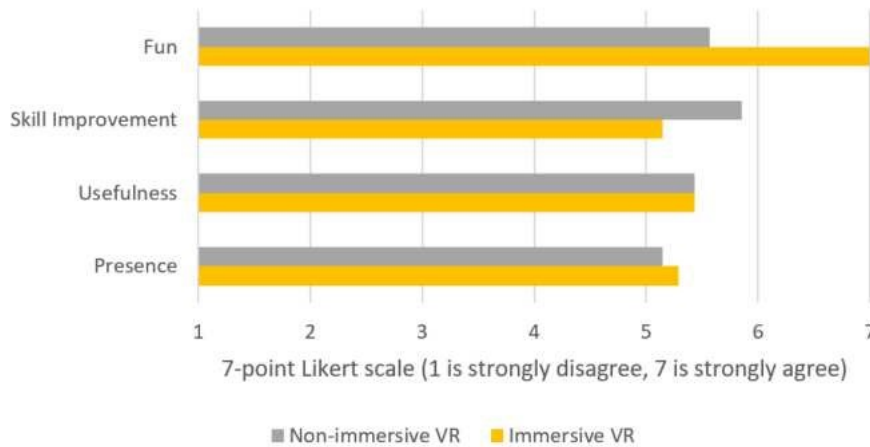


Abbildung 17: Ergebnisse zu Spaß, Fähigkeitsverbesserung, Nützlichkeit und Präsenz (Iteration 1)

Besonders auffällig war der Leistungsabfall zwischen Post- und Memory-Test, der statistisch signifikant war (*Wilcoxon-Test*, $p < .05$) und eine hohe Effektstärke aufwies (*Cohen's d* = 0.98). Dies deutet darauf hin, dass die Teilnehmer*innen nach der PC-Bedingung kurzfristig zwar lernten, das Gelernte jedoch über die Woche hinweg schlechter behielten.

Für die VR-Bedingung zeigten sich zwar ebenfalls Verbesserungen im Post-Test, doch diese waren statistisch nicht signifikant. Der Lernverlauf folgte demselben Muster wie bei der PC-Version (Anstieg im Post-Test, Rückgang im Memory-Test), allerdings auf einem etwas niedrigeren Gesamtniveau.

Beim Vergleich des relativen Wissenszuwachses (*Post minus Pre*) ergaben sich keine signifikanten Unterschiede zwischen den Bedingungen. Im Mittel lag der Zugewinn in der PC-Version bei 1.77 Punkten ($SD = 1.78$), in der VR-Version bei 1.00 Punkten ($SD = 1.94$). Auch der langfristige Wissenszuwachs (*Memory minus Pre*) war in beiden Fällen leicht negativ (*PC*: $M = -0.44$, *VR*: $M = -0.66$), was auf begrenzte Retention schließen lässt. Insgesamt zeigen die Ergebnisse, dass beide Varianten Lernfortschritte ermöglichen, aber die langfristige Wissenssicherung nicht ausreichend gegeben war.

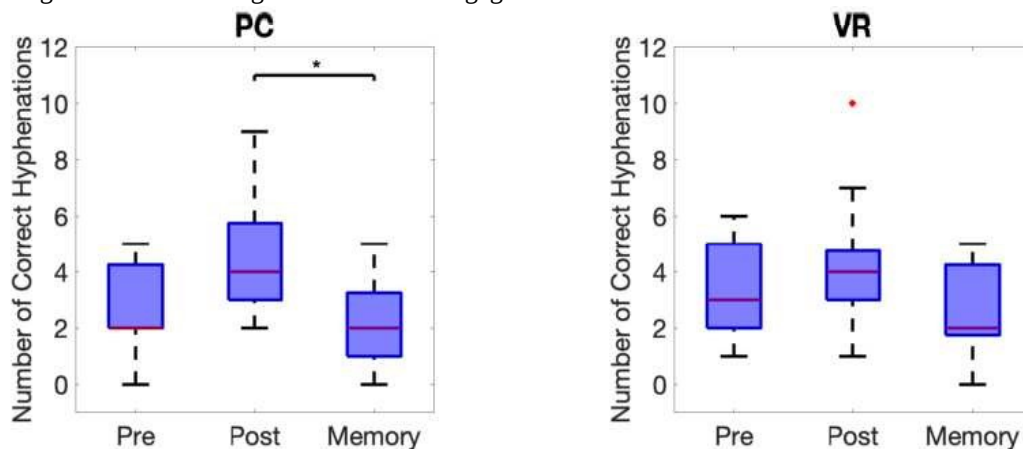


Abbildung 18: Ergebnisse des Pre-, Post- und Memoriesierungstests in PC und VR Variante.

In-Game-Leistung: Ein zentraler Aspekt der Untersuchung war das Spielverhalten der Teilnehmer*innen. Die Auswertung der Spielmetriken ergab, dass die Anzahl korrekt zusammengesetzter Silbenkombinationen in der PC-Version signifikant höher lag als in der VR-Version ($p < .01$, $d = 1.92$). Dies wird auch in Abbildung 19 veranschaulicht. Gleichzeitig benötigten die Teilnehmer*innen im VR-Modus deutlich mehr Zeit zwischen zwei richtigen Antworten ($p < .01$, $d = -1.37$). Die durchschnittliche Zeit pro richtiger Kombination war also in VR länger, was auf ein bedächtigeres Vorgehen schließen lässt (vgl. Abbildung 19).

Auffällig war auch die absolute Fehlerrate: In der PC-Version machten die Teilnehmer*innen signifikant mehr Fehler ($p < .01$, $d = 1.56$). Die relative Fehlerrate (Fehler pro richtiger Antwort) war ebenfalls höher in der PC-Version ($M = 1.18$, $SD = 0.42$) als in VR ($M = 0.95$, $SD = 0.47$), allerdings ohne statistische Signifikanz. Die Spielpunktzahl, welche auch das Besiegen von Gegnern berücksichtigte, lag in der PC-Version signifikant höher ($p < .01$, $d = 1.35$).

Diese Ergebnisse deuten auf einen interessanten Trade-off hin: In der PC-Version wurde effizienter gespielt, jedoch mit höherer Fehlerquote. In der VR-Version waren die Teilnehmer*innen langsamer, aber auch genauer. Unterstützt wird dies durch die Korrelationen: In der PC-Bedingung bestand eine starke positive Korrelation zwischen der Anzahl korrekter Kombinationen und der Anzahl der Fehler ($\rho = .74$, $p < .05$) - je mehr richtige Antworten, desto mehr Fehler. Zudem bestand eine starke negative Korrelation zwischen Fehleranzahl und Antwortzeit ($\rho = -.73$, $p < .05$), was für ein impulsives, wenig reflektiertes Spielverhalten spricht. In der VR-Bedingung traten diese Korrelationen nicht auf (ρ für Fehler und richtige Kombinationen: $-.15$, $p = .70$). Beide Bedingungen zeigten jedoch eine sehr starke negative Korrelation zwischen korrekten Kombinationen und Antwortzeit (je schneller, desto mehr richtige Antworten; PC: $\rho = -.99$, $p < .001$; VR: $\rho = -.99$, $p < .001$).

Diese Zusammenhänge deuten darauf hin, dass die Teilnehmer*innen in der PC-Version häufig eine „Trial-and-Error“-Strategie verfolgten (vgl. Abbildung 19), während sie in VR überlegter agierten. Die immersivere VR-Umgebung scheint somit gezieltere Entscheidungen zu fordern, vermutlich bedingt durch die stärkere körperliche Einbindung und die reduzierte Möglichkeit zum „schnellen Klicken“.

Usability, Flow und Spielspaß: Die wahrgenommene Usability beider Versionen wurde mit der System Usability Scale (SUS) erfasst. Die PC-Version wurde im Mittel mit 79.72 Punkten ($SD = 11.42$) bewertet, die R-Version mit 75.83 Punkten ($SD = 11.86$). Beide Werte liegen über der Schwelle für gute Usability (≥ 70), Unterschiede zwischen den Versionen waren jedoch nicht signifikant.

Auch das Flow-Erleben wurde in beiden Varianten als sehr hoch eingeschätzt ($M = 6.22$, $SD = 0.56$ auf einer 7-Punkte-Skala) - ein Hinweis auf eine starke Aufmerksamkeitsbindung und ein positives Spielerlebnis. Beim Item zur Spielfreude (*Das Spiel hat Spaß gemacht*) lagen die Werte nahezu identisch bei PC ($M = 6.89$, $SD = 0.33$) und VR ($M = 6.90$, $SD = 0.32$). Das Spiel wurde also unabhängig vom Modus als durchweg unterhaltsam empfunden.

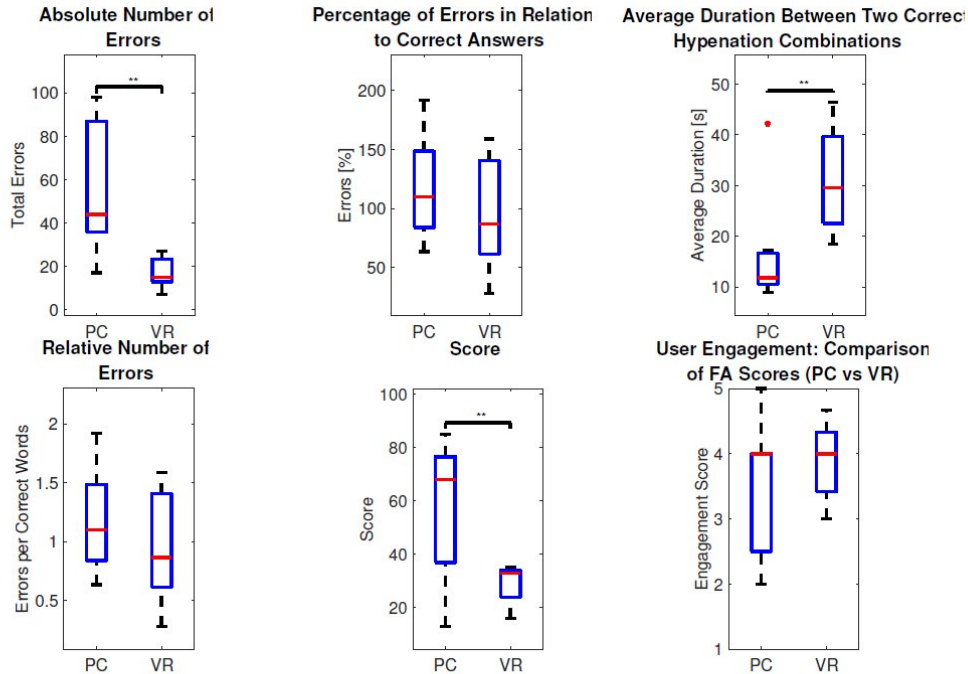


Abbildung 19: Boxplot-Ergebnisse für Spielmetriken: absolute Anzahl korrekter Silbenkombinationen, durchschnittliche Zeit zwischen korrekten Silbenkombinationen, absolute Anzahl Fehler, relative Anzahl Fehler, Punktzahl, und Subskala 'Focused Attention' der UES.

Engagement und Embodiment: Die User Engagement Scale (*kurzform*, 1-5) bestätigte ein insgesamt hohes Engagement in beiden Bedingungen (PC: $M = 4.31$, $SD = 0.44$; VR: $M = 4.44$, $SD = 0.39$). Besonders bei der Subskala „Fokussierte Aufmerksamkeit“ zeigte (vgl. Abbildung 19) sich eine leichte Tendenz zugunsten der VR-Version (PC: $M = 3.89$, $SD = 0.78$; VR: $M = 4.26$, $SD = 0.74$), wenn auch ohne statistische Signifikanz. Dies deutet jedoch darauf hin, dass VR eine intensivere kognitive Präsenz und Konzentration fördern könnte.

Das Gefühl der Körperzugehörigkeit in der VR-Version wurde ebenfalls erfasst (Embodiment). Die Teilnehmer*innen stimmten der Aussage „Ich hatte das Gefühl, die virtuelle Hand bzw. die virtuellen Füße gehörten zu meinem Körper“ stark zu ($M = 6.39$, $SD = 0.65$), was die immersive Qualität des Spiels unterstreicht und dessen potenziellen pädagogischen Wert durch körperlich verankertes Lernen bestätigt.

4.4 Konklusion

Zusammenfassend zeigte die Evaluation von „Syllable Soup“, dass ein SVRG zur Schriftsprachförderung bei gering literalisierten Erwachsenen prinzipiell wirksam und gut akzeptiert ist. Die Lernfortschritte waren in beiden Bedingungen vergleichbar, jedoch deuten Leistungsdaten und Korrelationen auf unterschiedliche Lernstrategien hin: Während am PC eher ausprobiert wurde, zeigte sich in VR ein fokussierteres, reflektierteres Lernverhalten. Dies wird durch höhere Aufmerksamkeitswerte und geringere Fehlerraten gestützt.

Besonders relevant ist die Erkenntnis, dass vollständig immersive VR nicht automatisch zu besseren Lernergebnissen führt, wohl aber zu veränderter Lernstrategie und höherer mentaler Präsenz. Diese Unterschiede sind für das Instructional Design, also der systematischen Planung, Entwicklung und Evaluation von Lerninhalten und -methoden, bedeutsam. VR scheint geeignet, Lernende starker zur bewussten Entscheidung zu führen, während niedrig-immersive Umgebungen schnelle, ggf. ungenauere Entscheidungen begünstigen.

Die Ergebnisse sprechen daher nicht für eine generelle Überlegenheit von VR, sondern für eine differenzierte Betrachtung: VR kann sinnvoll eingesetzt werden, wenn gezieltes, bewusstes Lernen im Vordergrund steht, insbesondere in sensiblen Bildungssettings mit Zielgruppen wie gering literalisierten Erwachsenen. Die Studie

hebt außerdem die Bedeutung von angepasster Interaktionsgestaltung, Feedbackmechanismen und kognitiver Entlastung hervor.

Fazit: Syllable Soup bietet eine vielversprechende, motivierende Lernumgebung für gering literalisierte Erwachsene. Zukünftige Arbeiten sollten die Ergebnisse mit größeren Stichproben und längeren Lernintervallen bestätigen sowie die Integration solcher Ansätze in formale Bildungsangebote prüfen.

5. Laut Manager - ein Desktop-basiertes Serious Game zur Schriftsprachförderung

5.1 Einführung



Abbildung 20: Laut-Manager Key Visual.

„Laut-Manager“ ist ein Serious Game für PCs und mobile Endgeräte (vgl. Abbildung 20), das ebenfalls im Rahmen von #ABCforJobs entwickelt wurde. Das Lernspiel dient der Verbesserung und Festigung orthographischen Wissens. Darüber hinaus fördert das Spiel phonologische Bewusstheit, Leseflüssigkeit sowie Merkfähigkeit von Spielenden im Zusammenhang mit Fachvokabular. Zielgruppe sind gering literalisierte Erwachsene aus der Hotel- und Gastronomiebranche.

„Laut-Manager“ greift dabei auf einer auf Punkten basierenden Spielmechanik zurück. Es gehört zum Genre der Sidescroller, einem Untergenre von Jump-’n’-Run- oder Rennspielen. Dieses zeichnet sich dadurch aus, dass die Spielfigur automatisch eine Strecke entlangläuft, ohne anzuhalten, während die Spieler*innen Hindernissen ausweichen und gegebenenfalls Gegenstände einsammeln müssen.

Die zugrundeliegenden Ziele des Spiels sind:

1. Das Ausweichen von „gegnerischen“ Objekten
2. Die Wahl des korrekten Lautes im Zusammenhang mit der gesuchten Lücke eines Wortes (beispielsweise: ss für das Wort Me__er - Messer)

Dabei steuern die Spielenden den Charakter „Staubi“ (vgl. Abbildung 21) durch einen endlos erscheinenden Hotelflur. Von Zeit zu Zeit hält das Spiel vor Zimmer- und Lifttüren an, aus welchen dann die zu wählenden Laute für das gesuchte Wort erscheinen. Der korrekte Laut muss daraufhin durch Klicken gewählt werden. Sofern der korrekte Laut gewählt wurde, erhalten die Spielenden einen Punkt und ein weiteres Lückenwort erscheint. Sobald eine falsche Auswahl getroffen wird, gibt das Spiel Rückmeldung, was die korrekte Antwort gewesen wäre, indem diese grün eingefärbt wird. Darüber hinaus muss Staubi bei seinen Säuberungsarbeiten

des Flures den fiesen Werkzeugen des Hausmeisterdienstes, dargestellt durch von rechts herannahende Hammer und Schraubenzieher, ausweichen. Staubli hat insgesamt drei Leben, die in Herzform dargestellt werden. Sind alle drei Leben verloren, fällt der Punktestand auf Null zurück. Gleichzeitig führt ein höherer Punktestand zu mehr und schneller entgegenkommenden Hausmeisterwerkzeugen. Die modulare Logik vom Laut-Manager, die auf eine Liste im JSON Format zurückgreift, erlaubt eine unkomplizierte Erweiterung oder Änderung der Lerninhalte. Das Spiel kann sowohl am PC wie auch auf mobilen Endgeräten wie Smartphones oder Tablets gespielt werden und bietet daher ein niederschwelliges Angebot für das Lernen „on the Go“. Die grafische Nutzerschnittstelle ist möglichst einfach gehalten, um der potenziellen Zielgruppe gering literalisierter Erwachsener gerecht zu werden.



Abbildung 21: Mockups des Charakters Staubli

5.2 Methodik

Zur Evaluation des Spiels wurde Gebrauch eines gemischten Methodenansatzes gemacht. Lernfortschritte wurden mittels Pre- und Post-Test ermittelt, hedonistische Qualität des Spiels mit Hilfe eines für die Zielgruppe reduzierten Fragebogens (vgl. Tabelle 1), der sich an standardisierten und validierten Fragebogen wie der Flow-Short Skala (FSS) [ER08] oder der General-Self-Efficacy Skala (GSE) orientiert. Die Leistung der Teilnehmer*innen wurde während des Spielens mit Hilfe der Fehlerquote sowie dem Punktestand gemessen.

Tabelle 1: Fragebogen zu Laut-Manager

Nr.	Item
1	Das Spiel ist nützlich.
2	Das Spiel unterstützt mich beim Lernen.
3	Das Spiel macht Spaß.
4	Das Spiel ist langweilig.
5	Das Spiel ist nicht zu schwer und nicht zu einfach.
6	Während dem Spielen habe ich die Zeit vergessen.
7	Ich war ganz und gar auf das Spiel konzentriert.
8	Das Spiel ist einfach zu bedienen.
9	Während dem Spiel habe ich mich geärgert.
10	Ich möchte das Spiel in der Zukunft benutzen.

Darüber hinaus wurden Fokusgruppendifkussionen, die einem halbstrukturierten Leitfaden folgten, durchgeführt (vgl. Tabelle 2).

Tabelle 2: Interviewleitfaden zu Laut-Manager

Frage	Item
1	Was hat Ihnen am Spiel gefallen und warum?
2	Was hat Ihnen am Spiel nicht gefallen und warum?
3	Wie haben Sie sich beim Spielen gefühlt und warum?
4	Gibt es etwas, das Sie am Spiel verändern würden, wenn Sie könnten und warum?
5	Gibt es sonst noch etwas, das Sie uns sagen mochten?

Insgesamt nahmen fünf Teilnehmer*innen an zwei Evaluationsterminen (*Oktober 2024 und Januar 2025*) teil. Von den Teilnehmer*innen gaben drei an weiblich zu sein sowie zwei an männlich zu sein. Der Altersdurchschnitt betrug 40.8 Jahre ($SD = 8.44$). Nach dem Pre-Test übten die Teilnehmer*innen mit Hilfe des Spiels für 15 Minuten (vgl. Abbildung 22). Zuvor hatten sie 3 Minuten Zeit, um sich an die Spielsteuerung zu gewöhnen. Anschließend folgte der Post-Test sowie der Fragebogen, zuletzt fand die Fokusgruppendifkussionen statt.

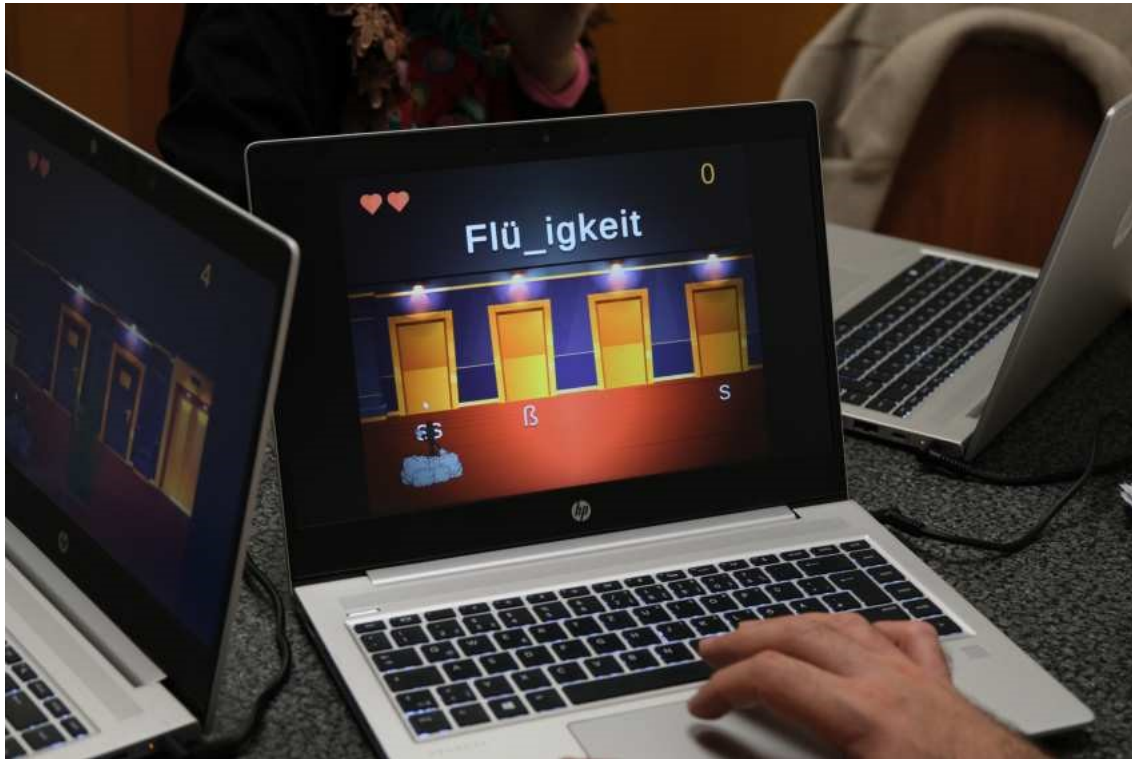


Abbildung 22: Teilnehmende spielen Laut-Manager.

5.3 Evaluationsergebnisse

Lernfortschritt: Bis auf T4, der sich sogar verschlechterte, verbesserten sich alle Teilnehmer*innen zwischen Pre- und Post-Test Ergebnissen, wenngleich dieser Unterschied statistisch nicht signifikant war (vgl. Abbildung 23). Insgesamt bearbeiteten die Teilnehmer*innen mehr Einzelaufgaben im Post-Test im Vergleich zum Pre-Test, was auf eine gewisse Verbesserung der Sicherheit in der Beantwortung schließen lässt. Prozentual lässt sich dieser Verbesserungstrend ebenfalls durch die anteilig beantworteten sowie anteilig korrekt beantworteten Fragen abbilden (vgl. Abbildung 24).

Hedonistische Qualität: Insgesamt wurde das Spiel als nützlich, effektiv, ausgewogen und unterhaltsam wahrgenommen, wenngleich diese Bewertungen im Widerspruch zu den Items Langeweile, Frustration und Benutzerfreundlichkeit stehen (vgl. Abbildung 25). Qualitative Ergebnisse der Fokusgruppendifkussionen relativieren jedoch die berichtete Langeweile im Fragebogen.

Leistung während des Spiels: Von den genutzten Wörtern aus dem Gesamtpool waren folgende Wörter besonders auffällig (vgl. Abbildung 26): Nussallergie, einschließen, Betriebsgröße, Stoßgefahr. Diese wurden sehr häufig falsch und erst spät korrekt gewählt. Wörter wie Gästebäder und Gießkanne wurden ausschließlich falsch angeklickt ($Error Rate = 1$). Diese Begriffe eignen sich als Kandidaten für verstärkte Übung oder didaktische Anpassungen im Spiel (z.B. deutlichere Hervorhebung der korrekten Schreibweise).

Im folgenden Liniendiagramm (vgl. Abbildung 27) wurde für jede korrekte Auswahl gezählt, wie viele Fehlklicks seit dem vorherigen Erfolg angefallen sind. So lassen sich verschiedene Phasen im Lernprozess ausmachen:

1. **Anfangsphase:** Bis zum 3. richtigen Wort waren häufig 3-10 Fehler nötig - der*die Spieler*in musste sich erst orientieren.
2. **Mid-Game-Plateau:** Danach sinken die Fehler pro Erfolg teilweise auf 0 (Index 5-9), was auf einen Lern- oder Gewöhnungseffekt hinweist.
3. **Späte Ausreißer:** Einzelne Spitzen (Index 10, 14, 20) deuten auf schwierige Wörter oder nachlassende Konzentration hin.

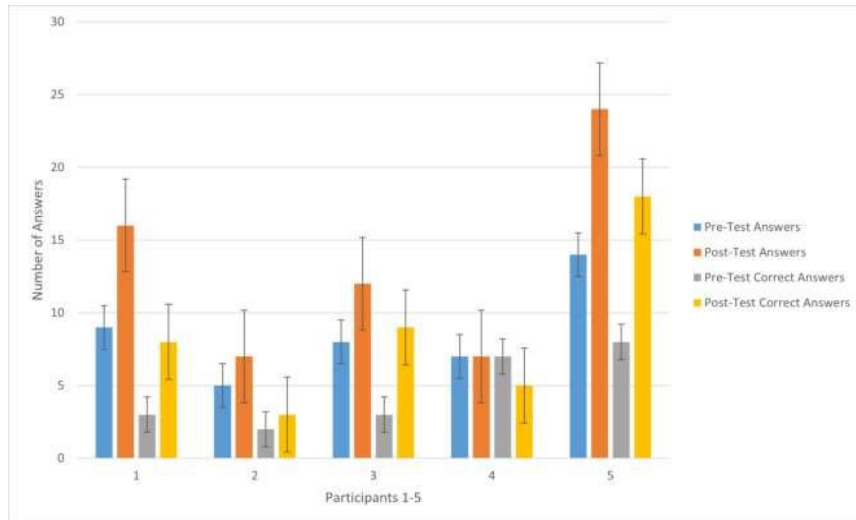


Abbildung 23: Anzahl der korrekten und falschen Antworten in Pre- und Post-Test.

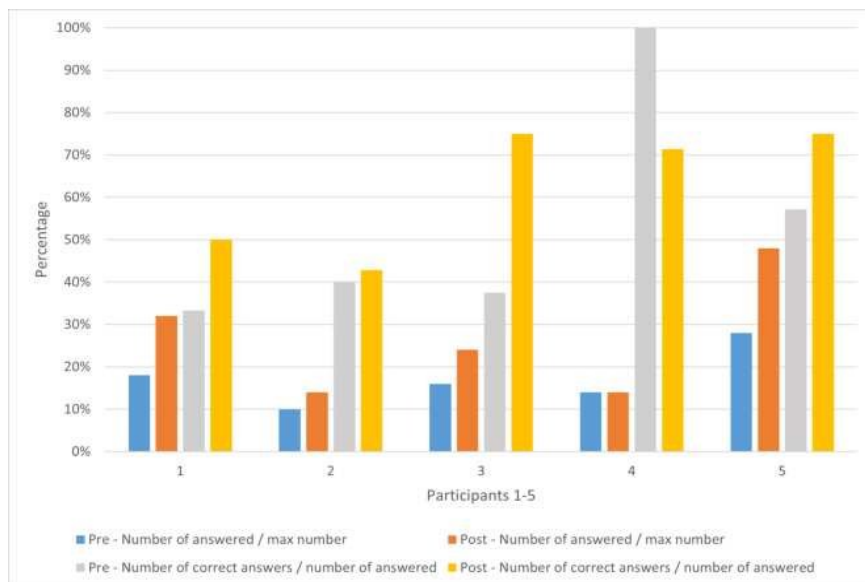


Abbildung 24: Anteil an korrekten und falschen Antworten im Verhältnis in Pre- und Post-Test.

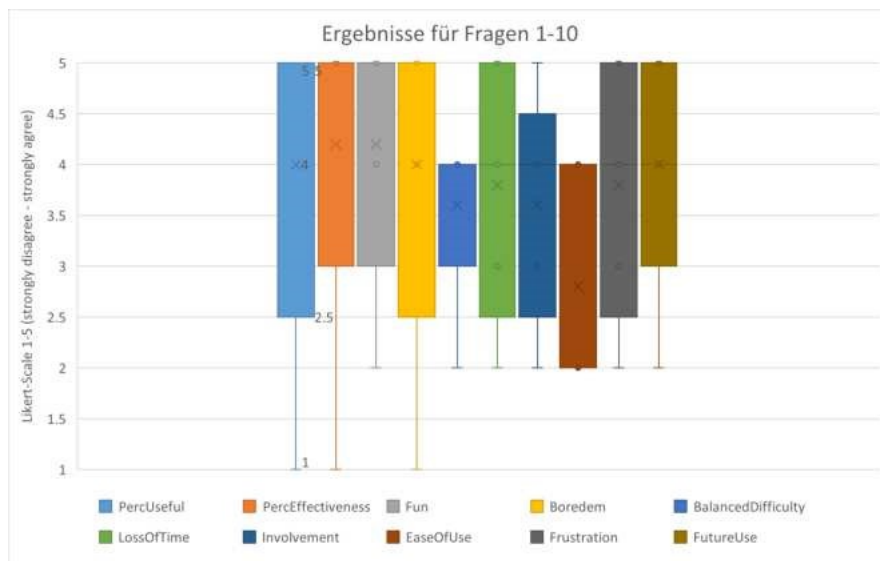


Abbildung 25: Antworten des Fragebogens als Boxplots.

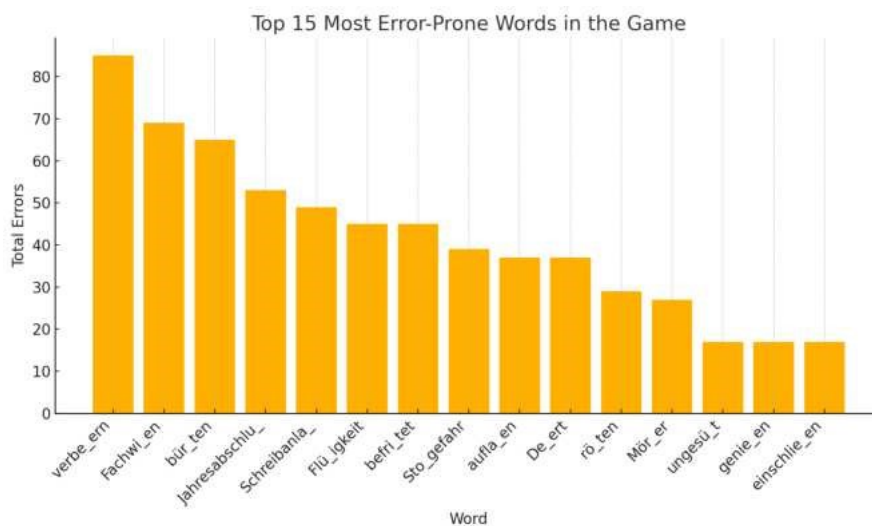


Abbildung 26: Das Schaubild zeigt die besonders fehleranfälligen Wörter.

In der Gesamtbilanz (Timestamp \approx 15:00) sind dies 20 richtige Wörter vs. 53 Fehlversuche (kumulative Error-Rate \approx 73%). Die hohe Gesamtfehlerrate spiegelt die bewusst knifflige Aufgabenstellung wider. Dennoch ist klar zu erkennen, dass die Fehlerrate innerhalb der Session schwankt und sich in Phasen deutlich verbessert, was ein positiver Hinweis auf Lernfortschritt darstellt.

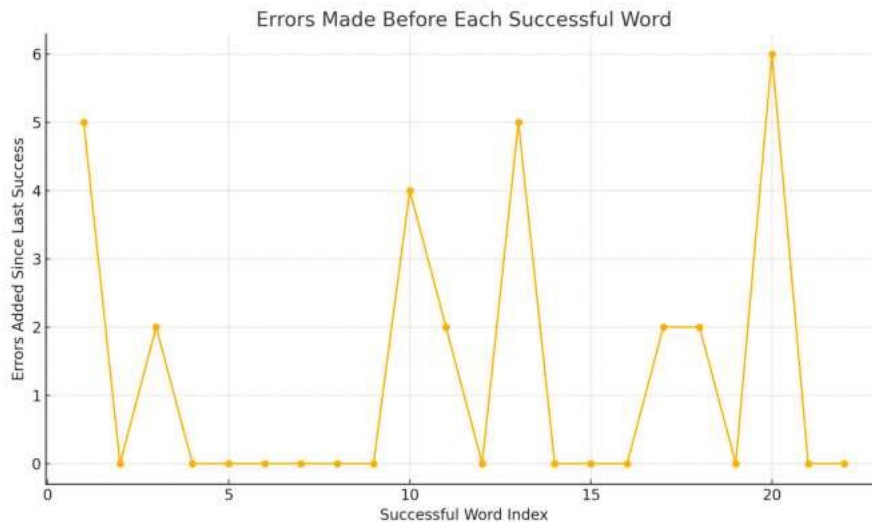


Abbildung 27: Das Schaubild zeigt die kumulative Fehlerentwicklung der Teilnehmer*innen während des Spiels.

Qualitative Ergebnisse der Fokusgruppendifkussionen: Bezüglich der Auswertung der Fokusgruppendifkussionen ergab eine induktive Kategorienbildung nach Mayring fünf Hauptkategorien:

1. **Wahrgenommener Lernnutzen:** Alle Teilnehmer*innen betonten, das Spiel sei „gut für Lernen“ und fördere den Erwerb neuer Wörter“. Die zweite Fokusgruppe ergänzte, sämtliche Begriffe seien neu und ließen sich memorieren“. Damit bestätigen die Daten die intendierte Wortschatzförderung.
2. **Motivation & Engagement:** Das Spiel wurde als „nicht langweilig“ und „interessant“ charakterisiert, korrekte Antworten erzeugten „Spaß“ und ein „Glücksgefühl“. Dies deutet auf hohes intrinsisches Engagement hin, das in der Gamification-Literatur als förderlich für Lernpersistenz gilt.
3. **Hemmende Spielmechaniken:** Gleichzeitig kritisierten die Teilnehmer*innen übereinstimmend die „Geschwindigkeit“ und mangelnde „Zeit zum Nachdenken“. Hinzu kamen eine „schwierige Steuerung“, was sich mit den Fragebogenergebnissen deckt, und übermäßige Hindernisse („viele Hammer“). Die Befunde illustrieren den Trade-off zwischen kognitiver Belastung und Flow-Erleben.
4. **emotionale Reaktionen:** Überwiegend wurden jedoch positive Emotionen („schon glücklich“) bei korrekten Lösungen berichtet. Negative Affekte traten bei Avatar-Verlust („Charakter tot“) oder dem als sehr schmerzhaft“ empfundenen Soundeffekt autsch“ auf. Das differenzierte Affektprofil unterstreicht die Bedeutung einer ausgewogenen Feedback-Gestaltung.
5. **Verbesserungsvorschläge:** Wiederkehrende Wünsche betrafen ein reduziertes Tempo bzw. eine Pausenfunktion, um mehr „Zeit zum Denken“ zu erhalten. Weitere Anregungen umfassten zusätzliche Lernspiele, Level-abhängige Hintergrundwechsel, Bild-Wort-Zuordnungen und die Abschwächung des Schmerz-Sounds. Diese Hinweise liefern konkrete Ansatzpunkte für weitere iterative Designzyklen.

5.4 Konklusion

Zusammenfassend zeigte die Evaluation von „Laut-Manager“, dass das Spiel seine didaktische Kernfunktion erfüllt, jedoch durch überhöhtes Tempo und sensorische Überlastung Frustrationspotenzial birgt. Eine adaptive Schwierigkeit, multimodale Feedbackoptionen und die Diversifizierung des Spielportfolios erscheinen geeignet, das positive Lern- und Emotionserleben noch nachhaltiger zu stabilisieren.

Darüber hinaus legen die vorliegenden Befunde nahe, dass „Laut-Manager“ bereits in seiner prototypischen Form einen messbaren, wenn auch noch geringen Lerngewinn erzeugt und gleichzeitig Freude am Üben

orthographischer Strukturen weckt. Um dieses Potenzial voll auszuschöpfen, könnten jedoch vier strategische Weiterentwicklungen entscheidend sein:

1. Feingranulare Adaptivität auf drei Ebenen

Tempo, *Aufgabenkomplexität* und *Ablenkungsgrad* sollten sich dynamisch an das aktuelle Kompetenz- und Stressniveau anpassen. Ein heuristischer Algorithmus konnte etwa nach fünf aufeinanderfolgenden Fehlklicks automatisch die Scrollgeschwindigkeit reduzieren, die Anzahl der Hindernisse halbieren und vorübergehend Hilfseinblendungen (silbengetrennte Wortvorschau) aktivieren. Die Evaluation zeigte, dass Fehlerrate und Frustration exakt in jenen Phasen anstiegen, in denen die kognitive Last Spitzenwerte erreichte. Mittels *Just-in-Time-Scaffolding* ließe sich der FlowKorridor stabilisieren, ohne den Spielcharakter aufzugeben.

2. Multimodale, wertschätzende Rückmeldungen

Negative Reize wie der als „schmerzhaft“ beschriebene Fehler-Sound sollten durch neutrale Fehlermarkierungen ersetzt werden, gefolgt von unmittelbar anschließender positiver Bestärkung „Du hast das Wort fast! Schau dir das -ss- an“. Kombiniert man visuelle Hervorhebung, Audiospur und gegebenenfalls haptisches Feedback (Vibration auf Mobilgeräten), wird die Fehlermarkierung verständlicher und weniger entmutigend — und adressiert damit die Diskrepanz zwischen Spaß und Frustration, die in den Fokusgruppen deutlich wurde.

3. Didaktische Erweiterung durch Mikro-Lernpfade

Die Analyse der Spieltelemetrie zeigte einzelne „Problemwörter“ mit Fehlerquoten von 100%. Für solche Items sollte das System automatisiert Kurzmodule nach dem Prinzip Spaced Repetition vorschlagen: Nach zwei Fehlversuchen wird das Wort als „schwierig“ markiert und taucht in den nächsten Sessions häufiger, jedoch in reduzierter Geschwindigkeit und unterstützt durch Bild- Wort-Zuordnung, auf. So lässt sich die geforderte Diversifizierung des Portfolios pragmatisch in den bestehenden Gameplay-Loop integrieren, ohne Medienbrüche zu erzeugen.

4. Blended-Learning-Szenario und Skalierung

Die sehr kleine Stichprobe limitiert die Verallgemeinerbarkeit. Künftige Untersuchungen sollten *Blended-Learning-Sets* erproben, in denen „Laut-Manager“ als Warm-up-Aktivität in Sprachkursen dient und Lehrkräfte über ein Dashboard Live-Daten (Fehlerraten, Problemwörter) erhalten. Das ermöglicht individuelle Nachbereitung und bindet das Spiel fest in institutionelle Lernpfade ein. Gleichzeitig erlaubt ein größeres N (> 50) quantitative Tests zu Effektstärken, Retention nach 4-8 Wochen und Transfer auf authentische Schreibaufgaben.

Parallel sollten Barrierefreiheits-Aspekte (Kontrast-Themes, Keyboard-Only-Steuerung) geprüft werden, um das Angebot auch für Lernende mit motorischen Einschränkungen zu öffnen und damit die angestrebte Niedrigschwelligkeit weiterzuentwickeln.

Ausblick: Wird „Laut-Manager“ unter diesen Prämissen weiterentwickelt, kann das Spiel mehr leisten als punktuelle Fehlerkorrekturen: Es könnte zum *situativen Worttrainer* werden, der orthographische Routine automatisiert, Selbstwirksamkeit stärkt und lernschwache Erwachsene spielerisch an das Schriftdeutsch heranführt. Die Kombination aus adaptiver Gamification, datengetriebener Didaktik und curricularer Verankerung adressiert genau jene Balance zwischen Motivation und Überforderung, die die Ergebnisse der ersten Evaluation als kritischen Erfolgsfaktor identifiziert haben. So ließe sich der aktuell noch fragile Lernerfolg in nachhaltigere Kompetenzentwicklung überführen und gleichzeitig ein beitragsfähiges Modell für andere Lernspiele im Bereich Grundbildung schaffen.

6. Zusammenfassung der Evaluationsergebnisse

Die Evaluation der vier Serious Games „Escape the Plate“, „Sky Dash“, „Syllable Soup“ und „Laut-Manager“ im Rahmen von #ABCforJobs verdeutlicht, dass spielbasierte digitale Lernformate ein hohes Potenzial zur Förderung schriftsprachlicher Kompetenzen bei gering literalisierten Erwachsenen besitzen. Die Ergebnisse zeigen, dass Serious Games, trotz unterschiedlicher Ausprägungen in Immersion, Komplexität und Interaktivität, in der Lage sind, Lernmotivation und Engagement zu steigern. Dies gilt insbesondere für Zielgruppen, die durch traditionelle Bildungsformate nur schwer erreicht werden.

Die Analysen bestätigen, dass spielerische Elemente wie unmittelbares Feedback, klare Zielstrukturen, adaptive Schwierigkeitsgrade und kontextnahe Lerninhalte zentrale Wirkmechanismen darstellen. Dabei erweisen sich Motivation und Spaß am Lernen als bedeutsame Faktoren für die Lernpersistenz, auch wenn die unmittelbare Wissenszuwachsrate nicht in allen Fällen über der klassischer Methoden liegt. Gleichzeitig wurde deutlich, dass kognitive Überlastung, zu hohe Spielgeschwindigkeit oder komplexe Steuerungselemente hemmend wirken können und durch gezieltes Game- und Instructional Design reduziert werden sollten.

Die vergleichende Betrachtung verschiedener Plattformen und Immersionsgrade, insbesondere bei „Syllable Soup“, zeigt, dass ein höheres Maß an Immersion nicht zwangsläufig zu besseren Lernergebnissen führt, wohl aber zu veränderter Lernstrategie und gesteigerter kognitiver Präsenz. Mobile und niedrighschwellige Zugänge, wie bei „Sky Dash“ und „Laut-Manager“, erweisen sich zudem als besonders praxistauglich, da sie flexible Lerngelegenheiten im Alltag eröffnen.

Für die Zielgruppe gering literalisierter Erwachsener liegt das besondere Potenzial von Serious Games in der Kombination aus motivationaler Ansprache, lebensweltlicher Relevanz und der Möglichkeit, Lerninhalte in kleinen, wiederholbaren Einheiten zu vermitteln. Durch iterative Weiterentwicklung, Anpassung an individuelle Lernvoraussetzungen sowie die Einbettung in hybride Lernarrangements kann dieses Potenzial weiter ausgeschöpft werden. Die vorliegenden Ergebnisse liefern damit nicht nur eine Grundlage für die Optimierung der getesteten Spiele, sondern auch mögliche Impulse für die zukünftige Gestaltung digitaler Bildungsangebote im Bereich der Grundbildung.

Literatur

- [BKM09] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.
- [CFSDT25] Inc Center For Self-Determination Theory. Intrinsic motivation inventory (imi), 2025. <https://selfdeterminationtheory.org/intrinsic-motivation-inventory/>.
- [ER08] Stefan Engeser and Falko Rheinberg. Flow, performance and moderators of challenge-skill balance. *Motivation and emotion*, 32:158–172, 2008.
- [OCH18] Heather L O’Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. *International Journal of Human-Computer Studies*, 112:28–39, 2018.

7 Anhang

7.1 IMI-Fragebogen in einfacher Sprache

Fragen zu Interesse/Freude	Fragen zu wahrgenommener Kompetenz	Fragen zu Wert/Nützlichkeit
1. Ich fand das Lernspiel wirklich toll.	1. Ich denke, ich war im Lernspiel ziemlich gut.	1. Ich denke, dieses Lernspiel könnte für mich nützlich sein.
2. Dieses Lernspiel hat mir Spaß gemacht.	2. Im Vergleich zu anderen Schülern denke ich, dass ich bei diesem Lernspiel ganz gut war.	2. Ich finde, dass es nützlich ist, dieses Lernspiel zu spielen.
3. Ich dachte, dieses Lernspiel war langweilig. (R)	3. Nachdem ich eine Weile das Lernspiel gespielt habe, fühlte ich mich ziemlich kompetent.	3. Ich denke, es ist wichtig, dieses Lernspiel zu spielen.
4. Das Lernspiel hat meine Aufmerksamkeit überhaupt nicht gefesselt. (R)	4. Ich bin mit meiner Leistung im Lernspiel zufrieden.	4. Ich wäre bereit, es wieder zu tun, weil es für mich etwas wertvoll ist.
5. Ich würde dieses Lernspiel als sehr interessant beschreiben.	5. Ich war ziemlich geschickt in diesem Lernspiel.	5. Ich denke, dass dieses Lernspiel mir helfen könnte.
6. Ich fand dieses Lernspiel ziemlich angenehm.	6. In diesem Lernspiel konnte ich nicht so gut sein. (R)	6. Ich glaube, dass es für mich von Nutzen sein könnte.
7. Während ich das Lernspiel gemacht habe, habe ich darüber nachgedacht, wie sehr ich es genossen habe.		7. Ich denke, dass dieses Lernspiel wichtig ist.

Abbildung 28: IMI-Fragebogen in einfacher Sprache

7.2 Fragebogen Sky Dash

No.	Item	Construct
1	The learning exercise was fun	Joy/UX
2	I know the 3-Letter-Codes	Knowledge/Self-Efficacy/UX
3	I think that I will be able to solve this task well in the future	Future Knowledge/Self-Efficacy/UX
4	I was fully concentrating on the task (game)	Focus/Flow/UX
5	I was aware of the progress I was making	Feedback/Flow/UX
6	I was challenged, but I could handle the challenge well	Challenge/Flow/UX
7	I thought about how long the task (game) would take	Loss of Time/Flow/UX

Abbildung 29: Fragebogen, der zur Erhebung der subjektiven Einschätzungen im Rahmen der „Sky Dash“ Studie eingesetzt wurde